(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 May 2002 (16.05.2002)

PCT

(10) International Publication Number
WO 02/38589 A2

(51) International Patent Classification[7]: C07K

(21) International Application Number: PCT/US01/47396

(22) International Filing Date:
9 November 2001 (09.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/247,943 9 November 2000 (09.11.2000) US

(71) Applicant *(for all designated States except US)*: **GENAISSANCE PHARMACEUTICALS, INC.** [US/US]; Five Science Park, New Haven, CT 06511 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ANASTASIO, Alison, E.** [US/US]; 367 Alden Avenue #8B, New Haven, CT 06515 (US). **CHEW, Anne** [US/US]; 1477 Beacon Street, #64, Brookline, MA 02446 (US). **CHOI, Julie, Y.** [US/US]; 38 Elizabeth Street, West Haven, CT 06516 (US). **DENTON, R., Rex** [US/US]; 129 Hunter's Trail, Madison, CT 06443 (US). **NANDABALAN, Krishnan** [US/US]; 228 Village Pond Road, Guilford, CT 06437 (US). **PETERSEN, Nathan** [US/US]; 300 Eastern Street, Apartment 288, New Haven, CT 06513 (US). **ROUNDS, Eileen** [US/US]; 36 Patton Drive, Cheshire, CT 06410 (US).

(74) Agents: **SHTIVELBAND, Inna** et al.; Genaissance Pharmaceuticals, Inc., Five Science Park, New Haven, CT 06511 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/38589 A2

(54) Title: HAPLOTYPES OF THE CYP2D6 GENE

(57) Abstract: Novel genetic variants of the Cytochrome P450, Subfamily IID, Polypeptide 6 (CYP2D6) gene are described. Various genotypes, haplotypes, and haplotype pairs that exist in the general United States population are disclosed for the CYP2D6 gene. Compositions and methods for haplotyping and/or genotyping the CYP2D6 gene in an individual are also disclosed. Polynucleotides defined by the haplotypes disclosed herein are also described

HAPLOTYPES OF THE CYP2D6 GENE

RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application Serial No. 60/247,943 filed November 9, 2000.

FIELD OF THE INVENTION

This invention relates to variation in genes that encode pharmaceutically-important proteins. In particular, this invention provides genetic variants of the human Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) gene and methods for identifying which variant(s) of this gene is/are possessed by an individual.

BACKGROUND OF THE INVENTION

Current methods for identifying pharmaceuticals to treat disease often start by identifying, cloning, and expressing an important target protein related to the disease. A determination of whether an agonist or antagonist is needed to produce an effect that may benefit a patient with the disease is then made. Then, vast numbers of compounds are screened against the target protein to find new potential drugs. The desired outcome of this process is a lead compound that is specific for the target, thereby reducing the incidence of the undesired side effects usually caused by activity at non-intended targets. The lead compound identified in this screening process then undergoes further *in vitro* and *in vivo* testing to determine its absorption, disposition, metabolism and toxicological profiles. Typically, this testing involves use of cell lines and animal models with limited, if any, genetic diversity.

What this approach fails to consider, however, is that natural genetic variability exists between individuals in any and every population with respect to pharmaceutically-important proteins, including the protein targets of candidate drugs, the enzymes that metabolize these drugs and the proteins whose activity is modulated by such drug targets. Subtle alteration(s) in the primary nucleotide sequence of a gene encoding a pharmaceutically-important protein may be manifested as significant variation in expression, structure and/or function of the protein. Such alterations may explain the relatively high degree of uncertainty inherent in the treatment of individuals with a drug whose design is based upon a single representative example of the target or enzyme(s) involved in metabolizing the drug. For example, it is well-established that some drugs frequently have lower efficacy in some individuals than others, which means such individuals and their physicians must weigh the possible benefit of a larger dosage against a greater risk of side effects. Also, there is significant variation in how well people metabolize drugs and other exogenous chemicals, resulting in substantial interindividual variation in the toxicity and/or efficacy of such exogenous substances (Evans et al., 1999, *Science* 286:487-491). This variability in efficacy or toxicity of a drug in genetically-diverse patients makes many drugs ineffective or even dangerous in certain groups of the population, leading to the failure of

such drugs in clinical trials or their early withdrawal from the market even though they could be highly beneficial for other groups in the population. This problem significantly increases the time and cost of drug discovery and development, which is a matter of great public concern.

It is well-recognized by pharmaceutical scientists that considering the impact of the genetic variability of pharmaceutically-important proteins in the early phases of drug discovery and development is likely to reduce the failure rate of candidate and approved drugs (Marshall A 1997 *Nature Biotech* **15**:1249-52; Kleyn PW et al. 1998 *Science* **281**: 1820-21; Kola I 1999 *Curr Opin Biotech* **10**:589-92; Hill AVS et al. 1999 in *Evolution in Health and Disease* Stearns SS (Ed.) Oxford University Press, New York, pp 62-76; Meyer U.A. 1999 in *Evolution in Health and Disease* Stearns SS (Ed.) Oxford University Press, New York, pp 41-49; Kalow W et al. 1999 *Clin. Pharm. Therap.* **66**:445-7; Marshall, E 1999 *Science* **284**:406-7; Judson R et al. 2000 *Pharmacogenomics* **1**:1-12; Roses AD 2000 *Nature* **405**:857-65). However, in practice this has been difficult to do, in large part because of the time and cost required for discovering the amount of genetic variation that exists in the population (Chakravarti A 1998 *Nature Genet* **19**:216-7; Wang DG et al 1998 *Science* **280**:1077-82; Chakravarti A 1999 *Nat Genet* **21**:56-60 (suppl); Stephens JC 1999 *Mol. Diagnosis* **4**:309-317; Kwok PY and Gu S 1999 *Mol. Med. Today* **5**:538-43; Davidson S 2000 *Nature Biotech* **18**:1134-5).

The standard for measuring genetic variation among individuals is the haplotype, which is the ordered combination of polymorphisms in the sequence of each form of a gene that exists in the population. Because haplotypes represent the variation across each form of a gene, they provide a more accurate and reliable measurement of genetic variation than individual polymorphisms. For example, while specific variations in gene sequences have been associated with a particular phenotype such as disease susceptibility (Roses AD *supra*; Ulbrecht M et al. 2000 *Am J Respir Crit Care Med* **161**: 469-74) and drug response (Wolfe CR et al. 2000 *BMJ* **320**:987-90; Dahl BS 1997 *Acta Psychiatr Scand* **96** (Suppl 391): 14-21), in many other cases an individual polymorphism may be found in a variety of genomic backgrounds, i.e., different haplotypes, and therefore shows no definitive coupling between the polymorphism and the causative site for the phenotype (Clark AG et al. 1998 *Am J Hum Genet* **63**:595-612; Ulbrecht M et al. 2000 *supra*; Drysdale et al. 2000 *PNAS* **97**:10483-10488). Thus, there is an unmet need in the pharmaceutical industry for information on what haplotypes exist in the population for pharmaceutically-important genes. Such haplotype information would be useful in improving the efficiency and output of several steps in the drug discovery and development process, including target validation, identifying lead compounds, and early phase clinical trials (Marshall et al., *supra*).

One pharmaceutically-important gene for the treatment of hypertension, atrial and ventricular arrhythmias, Parkinson's disease and drug-induced lupus syndrome is the Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) gene or its encoded product. CYP2D6 is an enzyme that belongs to the cytochrome P450 family, whose members are responsible for the detoxification of many drugs and environmental chemicals. CYP2D6 is involved in the metabolism of drugs such as

antiarythmics, adrenoceptor antagonists and trycyclic antidepressants (SWISS-PROT:P10635).

CYP2D6 was originally identified as the primary enzyme responsible for the hydroxylation of the hypertension drug, debrisoquine to its metabolite, 4-hydroxydebrisoquine. Clinically, patients exhibit wide variation in hypotensive response to treatment with debrisoquine, which has been attributed to polymorphisms in the CYP2D6 gene that alter the enzymes ability to metaboloize, ie, hydroxylate, this drug (Mahgoub et al., 1977, *Lancet* 2: 584-586). Clinically significant inherited variations of CYP2D6-mediated drug metabolism are characterized by two phenotypes: the extensive metabolizer (EM) and the poor metabolizer (PM). 5-10% of individuals in Caucasian populations are of the PM phenotype and have deficient metabolism of debrisoquine and over 25 other drugs (Kagimoto et al., *J Biol Chem* 1990 Oct 5;265(28):17209-14). Lennard et al. (*N. Engl. J Med* 1982; 307:1558-1560) demonstrated that metabolism of metoprolol, a beta-1-selective adrenoceptor antagonist used for the treatment of angina, may be affected by CYP2D6 polymorphisms. In poor hydroxylators, a single daily dose of metoprolol may control angina, whereas in 'extensive hydroxylators,' 2 or 3 doses a day may be necessary, since plasma metoprolol concentrations may remain negligible 24 hours after dosing.

The antiarrhythmic agent propafenone, also metabolized by CYP2D6, is effective in the management of atrial and ventricular arrhythmias and has been shown to have variable efficacy as a beta-blocker (Lennard et al., 1983, *Pharm. Int.* 4: 61-65). Lee et al. (*N. Engl. J Med* 1990; 322:1764-1768) presented evidence that genetically determined variations in the conversion of propafenone to its 5-hydroxy metabolite accounts for variations in the drug's beta-blocking action. Slow metabolizers showed greater beta-blockade than rapid metabolizers did at a lower drug dosage, with comparable responses seen with the higher doses.

The pathogenesis of Parkinson's disease may be influenced by genetic and environmental factors. Cytochrome P450 mono-oxygenases, such as CYP2D6, help to protect against toxic environmental compounds and individual variations in cytochrome P450 expression might, therefore, influence susceptibility to environmentally linked diseases (Smith et al., *Lancet* 1992; 339:1375-1377). Studies show that significantly more parkinsonian than control subjects had partially or totally defective 4-hydroxylation of debrisoquine due to polymorphisms in the CYP2D6 (Barbeau et al., *Lancet* 1985; 2:1213-1216). The authors indicate that poor metabolizers of debrisoquine tended to have had earlier onset of disease.

CYP2D6 is also the major isozyme involved in the formation of N-hydroxyprocainamide, a metabolite potentially involved in the drug-induced lupus syndrome observed with procainamide (Lessard et al., *Pharmacogenetics* 1997; 7:381-390). The occurrence of lupus-like syndrome in a significant number of patients treated with procainamide has limited the clinical use of this antiarrhythmic drug. Further studies may show that a low CYP2D6 activity, either genetically determined or pharmacologically modulated, could prevent drug-induced lupus syndrome observed during chronic therapy with procainamide (Lessard et al., *Pharmacogenetics* 1999; 9:683-696).

The Cytochrome P450, subfamily IID, Polypeptide 6 gene is located on chromosome 22q13.1 and contains 9 exons that encode a 497 amino acid protein. A reference sequence for the CYP2D6 gene is shown in the contiguous lines of Figure 1 (Genaissance Reference No. 3578126; SEQ ID NO: 1). Reference sequences for the coding sequence (GenBank Accession No. NM_000106.2) and protein are shown in Figures 2 (SEQ ID NO: 2) and 3 (SEQ ID NO: 3), respectively.

Thirteen polymorphisms have been reported in the CYP2D6 gene which are refered to herein as PS5, PS7, PS8, PS9, PS11, PS13, PS21, PS25, PS30, PS31, PS33, PS35, and PS41.

A polymorphism of guanine or adenine (PS5) at a position corresponding to nucleotide position 825 in Figure 1 has been reported in the NCBI SNP Database (rs#1080993). A polymorphism of guanine or adenine (PS7) at a position corresponding to nucleotide position 1019 in Figure 1 which resutlts in the amino acid varaition of a valine or methionine at amino acid postion 7 in Figure 3 has been reported in the literature (Marez et al., *Pharmacogenetics* 1997; 7:193-202). A polymorphism of guanine or adenine (PS8) at a position corresponding to nucleotide position 1031 in Figure 1 that results in an amino acid variation of valine or methionine at a position corresponding to amino acid 11 has been identified (SWISS-PROT: P10635).

Kagimoto et al., (*supra*) identified a polymorphism of cytosine or thymine (PS9) at a position corresponding to nucleotide 1100 in Figure 1, which results in a variation of proline or serine at a position corresponding to amino acid 34 in Figure 3. A polymorphism of thymine or guanine (PS11) at a position corresponding to nucleotide position 1843 in Figure 1 has been reported in the NCBI SNP Database (rs# 769261). A polymorphism of a cytosine or adenine (PS13) and a polymorphism of an adenine or guanine (PS14) at nucleotide positions 1974 and 1984, respectively in Figure 1 have been reported in the literature (Marez et al., supra). These polymorphisms result in amino acid variations of leucine or methionine and histidine or arginine at amino acid positions 91 and 94, respectively in Figure 3.

A polymorphism of cytosine or thymine (PS21) at a position corresponding to nucleotide position 2039 in Figure 1 (NCBI SNP Database rs# 1081003) and a polymorphism of guanine or adenine (PS25) at a position corresponding to nucleotide position 2170 in Figure 1 (NCBI SNP Databse rs#1081004) have been reported. A polymorphism of guanine ar cytosine at a position corresponding to nucleotide position 2661 in Figure 1 (PS30) has been reported in the NCBI SNP Database (rs# 1058164). This polymorphism results in a valine or isoleucine at a position corresponding to amino acid position 136 in Figure 3.

A polymorphism of cytosine or guanine (PS31) at a position corresponding to nucleotide position 2704 in Figure 1 has been reported in literature. This polymorphism results in the amin acid variation of of a glutamine or glutamate at amino acid position 151 in Figure 3 (Marez et al., *supra*). A polymorphisms of a guanine or adenine (PS33) at a position corresponding to nucleotide position 2846 in Figure has been reproted in the literature (Kagimoto et al., *supra*).

A polymorphism of a thymine or cytosine (PS35) at a position corresponding to nucleotide

position 3470 in Figure 1 has been reported in the NCBI SNP Database (rs# 1058169). A polymorphism of guanine or cytosine (PS41) at a position corresponding to nucleotide position 5180 in Figure 1, which causes an amino acid variation of serine or threonine at a position corresponding to amino acid 486 in Figure 3, was reported by (Yokota et al., *Pharmacogenetics* 1993; 3:256-263).

Because of the potential for variation in the CYP2D6 gene to affect the expression and function of the encoded protein, it would be useful to know whether additional polymorphisms exist in the CYP2D6 gene, as well as how such polymorphisms are combined in different copies of the gene. Such information could be applied for studying the biological function of CYP2D6 as well as in identifying drugs targeting this protein for the treatment of disorders related to its abnormal expression or function.

SUMMARY OF THE INVENTION

Accordingly, the inventors herein have discovered 29 novel polymorphic sites in the CYP2D6 gene. These polymorphic sites (PS) correspond to the following nucleotide positions in Figure 1: 636 (PS1), 678 (PS2), 769 (PS3), 776 (PS4), 915 (PS6), 1827 (PS10), 1966 (PS12), 1984 (PS14), 1997 (PS15), 2014 (PS16), 2022 (PS17), 2023 (PS18), 2028 (PS19), 2036 (PS20), 2062 (PS22), 2067 (PS23), 2118 (PS24), 2179 (PS26), 2611 (PS27), 2635 (PS28), 2659 (PS29), 2716 (PS32), 3292 (PS34), 4183 (PS36), 4201 (PS37), 4254 (PS38), 4384 (PS39), 4435 (PS40) and 5212 (PS42). The polymorphisms at these sites are guanine or adenine at PS1, thymine or cytosine at PS2, guanine or cytosine at PS3, adenine or guanine at PS4, thymine or cytosine at PS6, guanine or cytosine at PS10, guanine or adenine at PS12, adenine or guanine at PS14, cytosine or guanine at PS15, thymine or cytosine at PS16, adenine or thymine at PS17, cytosine or thymine at PS18, adenine or guanine at PS19, thymine or cytosine at PS20, adenine or guanine at PS22, thymine or guanine at PS23, cytosine or thymine at PS24, guanine or cytosine at PS26, thymine or adenine at PS27, thymine or cytosine at PS28, guanine or adenine at PS29, guanine or adenine at PS32, guanine or adenine at PS34, guanine or adenine at PS36, cytosine or thymine at PS37, thymine or cytosine at PS38, adenine or cytosine at PS39, cytosine or adenine at PS40 and cytosine or thymine at PS42. In addition, the inventors have determined the identity of the alleles at these sites, as well as at the previously identified sites at nucleotide positions 825 (PS5), 1019 (PS7), 1031 (PS8), 1100 (PS9), 1843 (PS11), 1974 (PS13), 2039 (PS21), 2170 (PS25), 2661 (PS30), 2704 (PS31), 2846 (PS33), 3470 (PS35) and 5180 (PS41), in a human reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: African descent, Asian, Caucasian and Hispanic/Latino. From this information, the inventors deduced a set of haplotypes and haplotype pairs for PS1-PS42 in the CYP2D6 gene, which are shown below in Tables 4 and 3, respectively. Each of these CYP2D6 haplotypes constitutes a code that defines the variant nucleotides that exist in the human population at this set of polymorphic sites in the CYP2D6 gene. Thus each CYP2D6 haplotype also represents a naturally-occurring isoform (also referred to herein as an "isogene") of the CYP2D6 gene. The

frequency of each haplotype and haplotype pair within the total reference population and within each of the four major population groups included in the reference population was also determined.

Thus, in one embodiment, the invention provides a method, composition and kit for genotyping the CYP2D6 gene in an individual. The genotyping method comprises identifying the nucleotide pair that is present at one or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS6, PS10, PS12, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS22, PS23, PS24, PS26, PS27, PS28, PS29, PS32, PS34, PS36, PS37, PS38, PS39, PS40 and PS42 in both copies of the CYP2D6 gene from the individual. A genotyping composition of the invention comprises an oligonucleotide probe or primer which is designed to specifically hybridize to a target region containing, or adjacent to, one of these novel CYP2D6 polymorphic sites. A genotyping kit of the invention comprises a set of oligonucleotides designed to genotype each of these novel CYP2D6 polymorphic sites. In a preferred embodiment, the genotyping kit comprises a set of oligonucleotides designed to genotype each of PS1-PS42. The genotyping method, composition, and kit are useful in determining whether an individual has one of the haplotypes in Table 4 below or has one of the haplotype pairs in Table 3 below.

The invention also provides a method for haplotyping the CYP2D6 gene in an individual. In one embodiment, the haplotyping method comprises determining, for one copy of the CYP2D6 gene, the identity of the nucleotide at one or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS6, PS10, PS12, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS22, PS23, PS24, PS26, PS27, PS28, PS29, PS32, PS34, PS36, PS37, PS38, PS39, PS40 and PS42. In another embodiment, the haplotyping method comprises determining whether one copy of the individual's CYP2D6 gene is defined by one of the CYP2D6 haplotypes shown in Table 4, below, or a sub-haplotype thereof. In a preferred embodiment, the haplotyping method comprises determining whether both copies of the individual's CYP2D6 gene are defined by one of the CYP2D6 haplotype pairs shown in Table 3 below, or a sub-haplotype pair thereof. Establishing the CYP2D6 haplotype or haplotype pair of an individual is useful for improving the efficiency and reliability of several steps in the discovery and development of drugs for treating diseases associated with CYP2D6 activity, e.g., hypertension, atrial and ventricular arrhythmias Parkinson's disease and drug-induced lupus syndrome.

For example, the haplotyping method can be used by the pharmaceutical research scientist to validate CYP2D6 as a candidate target for treating a specific condition or disease predicted to be associated with CYP2D6 activity. Determining for a particular population the frequency of one or more of the individual CYP2D6 haplotypes or haplotype pairs described herein will facilitate a decision on whether to pursue CYP2D6 as a target for treating the specific disease of interest. In particular, if variable CYP2D6 activity is associated with the disease, then one or more CYP2D6 haplotypes or haplotype pairs will be found at a higher frequency in disease cohorts than in appropriately genetically matched controls. Conversely, if each of the observed CYP2D6 haplotypes are of similar frequencies in the disease and control groups, then it may be inferred that variable

CYP2D6 activity has little, if any, involvement with that disease. In either case, the pharmaceutical research scientist can, without *a priori* knowledge as to the phenotypic effect of any CYP2D6 haplotype or haplotype pair, apply the information derived from detecting CYP2D6 haplotypes in an individual to decide whether modulating CYP2D6 activity would be useful in treating the disease.

The claimed invention is also useful in screening for compounds targeting CYP2D6 to treat a specific condition or disease predicted to be associated with CYP2D6 activity. For example, detecting which of the CYP2D6 haplotypes or haplotype pairs disclosed herein are present in individual members of a population with the specific disease of interest enables the pharmaceutical scientist to screen for a compound(s) that displays the highest desired agonist or antagonist activity for each of the CYP2D6 isoforms present in the disease population, or for only the most frequent CYP2D6 isoforms present in the disease population. Thus, without requiring any *a priori* knowledge of the phenotypic effect of any particular CYP2D6 haplotype or haplotype pair, the claimed haplotyping method provides the scientist with a tool to identify lead compounds that are more likely to show efficacy in clinical trials.

Haplotyping the CYP2D6 gene in an individual is also useful to control for genetically-based bias in the design of candidate drugs that target or are metabolized by CYP2D6. For example, for a lead compound that is metabolized by CYP2D6, the pharmaceutical scientist of ordinary skill would be concerned that a favorable efficacy and/or side effect profile shown in a Phase II or Phase III trial may not be replicated in the general population if a higher (or lower) percentage of patients in the treatment group, compared to the general population, have a form of the CYP2D6 gene that makes them genetically predisposed to metabolize the drug more efficiently than patients with other forms of the CYP2D6 gene. Similarly, this pharmaceutical scientist would recognize the potential for bias in the results of a Phase II or Phase III clinical trial of a drug targeting CYP2D6 that could be introduced if individuals whose CYP2D6 gene structure makes them genetically predisposed to respond well to the drug are present in a higher (or lower) frequency in the treatment group than in the control group (Bacanu et al., 2000, *Am. J. Hum. Gen.* **66**:1933-44; Pritchard et al., 2000, *Am. J. Hum. Gen.* **67**: 170-81).

The pharmaceutical scientist can immediately reduce this potential for genetically-base bias in the results of clinical trials of drugs metabolized by or targeting CYP2D6 by practicing the claimed invention. In particular, by determining which of the CYP2D6 haplotypes disclosed herein are present in individuals recruited to participate in a clinical trial of a drug metabolized by or targeting CYP2D6, the pharmaceutical scientist can then assign that individual to the treatment or control group as appropriate to ensure that approximately equal frequencies of different CYP2D6 haplotypes (or haplotype pairs) are represented in the two groups and/or the frequencies of different CYP2D6 haplotypes or haplotype pairs are similar to the frequencies in the general population. Thus, by practicing the claimed invention, the pharmaceutical scientist can more confidently rely on the information learned from the trial, without first determining the phenotypic effect of any CYP2D6

haplotype or haplotype pair.

In another embodiment, the invention provides a method for identifying an association between a trait and a CYP2D6 genotype, haplotype, or haplotype pair for one or more of the novel polymorphic sites described herein. The method comprises comparing the frequency of the CYP2D6 genotype, haplotype, or haplotype pair in a population exhibiting the trait with the frequency of the CYP2D6 genotype or haplotype in a reference population. A higher frequency of the CYP2D6 genotype, haplotype, or haplotype pair in the trait population than in the reference population indicates the trait is associated with the CYP2D6 genotype, haplotype, or haplotype pair. In preferred embodiments, the trait is susceptibility to a disease, severity of a disease, the staging of a disease or response to a drug. In a particularly preferred embodiment, the CYP2D6 haplotype is selected from the haplotypes shown in Table 4, or a sub-haplotype thereof. Such methods have applicability in developing diagnostic tests and therapeutic treatments for hypertension, atrial and ventricular arrhythmias, Parkinson's disease and drug-induced lupus syndrome.

In yet another embodiment, the invention provides an isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the CYP2D6 gene or a fragment thereof. The reference sequence comprises the contiguous sequences shown in Figure 1 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of adenine at PS1, cytosine at PS2, cytosine at PS3, guanine at PS4, cytosine at PS6, cytosine at PS10, adenine at PS12, guanine at PS14, guanine at PS15, cytosine at PS16, thymine at PS17, thymine at PS18, guanine at PS19, cytosine at PS20, guanine at PS22, guanine at PS23, thymine at PS24, cytosine at PS26, adenine at PS27, cytosine at PS28, adenine at PS29, adenine at PS32, adenine at PS34, adenine at PS36, thymine at PS37, cytosine at PS38, cytosine at PS39, adenine at PS40 and thymine at PS42. In a preferred embodiment, the polymorphic variant comprises one or more additional polymorphisms selected from the group consisting of adenine at PS5, adenine at PS7, adenine at PS8, thymine at PS9, guanine at PS11, adenine at PS13, thymine at PS21, adenine at PS25, cytosine at PS30, guanine at PS31, adenine at PS33, cytosine at PS35 and cytosine at PS41.

A particularly preferred polymorphic variant is an isogene of the CYP2D6 gene. A CYP2D6 isogene of the invention comprises guanine or adenine at PS1, thymine or cytosine at PS2, guanine or cytosine at PS3, adenine or guanine at PS4, guanine or adenine at PS5, thymine or cytosine at PS6, guanine or adenine at PS7, guanine or adenine at PS8, cytosine or thymine at PS9, guanine. The invention also provides a collection of CYP2D6 isogenes, referred to herein as a CYP2D6 genome anthology.

In another embodiment, the invention provides a polynucleotide comprising a polymorphic variant of a reference sequence for a CYP2D6 cDNA or a fragment thereof. The reference sequence comprises SEQ ID NO:2 (Fig.2) and the polymorphic cDNA comprises at least one polymorphism selected from the group consisting of adenine at a position corresponding to nucleotide 263, guanine at a position corresponding to nucleotide 281, guanine at a position corresponding to nucleotide 294,

cytosine at a position corresponding to nucleotide 311, thymine at a position corresponding to nucleotide 319, thymine at a position corresponding to nucleotide 320, guanine at a position corresponding to nucleotide 325, cytosine at a position corresponding to nucleotide 333, adenine at a position corresponding to nucleotide 358, cytosine at a position corresponding to nucleotide 382, adenine at a position corresponding to nucleotide 406, adenine at a position corresponding to nucleotide 463, adenine at a position corresponding to nucleotide 1012, thymine at a position corresponding to nucleotide 1030, cytosine at a position corresponding to nucleotide 1083 and thymine at a position corresponding to nucleotide 1489. In a preferred embodiment, the polymorphic variant comprises one or more additional polymorphisms selected from the group consisting of adenine at a position corresponding to nucleotide 19, adenine at a position corresponding to nucleotide 31, thymine at a position corresponding to nucleotide 100, adenine at a position corresponding to nucleotide 271, thymine at a position corresponding to nucleotide 336, cytosine at a position corresponding to nucleotide 408, guanine at a position corresponding to nucleotide 451, cytosine at a position corresponding to nucleotide 696 and cytosine at a position corresponding to nucleotide 1457. A particularly preferred polymorphic cDNA variant comprises the coding sequence of a CYP2D6 isogene defined by haplotypes 1-10, 13-16, 18, 19, 21, 22, and 24-34.

Polynucleotides complementary to these CYP2D6 genomic and cDNA variants are also provided by the invention. It is believed that polymorphic variants of the CYP2D6 gene will be useful in studying the expression and function of CYP2D6, and in expressing CYP2D6 protein for use in screening for candidate drugs to treat diseases related to CYP2D6 activity.

In other embodiments, the invention provides a recombinant expression vector comprising one of the polymorphic genomic and cDNA variants operably linked to expression regulatory elements as well as a recombinant host cell transformed or transfected with the expression vector. The recombinant vector and host cell may be used to express CYP2D6 for protein structure analysis and drug binding studies.

In yet another embodiment, the invention provides a polypeptide comprising a polymorphic variant of a reference amino acid sequence for the CYP2D6 protein. The reference amino acid sequence comprises SEQ ID NO:3 (Fig.3) and the polymorphic variant comprises at least one variant amino acid selected from the group consisting of histidine at a position corresponding to amino acid position 88, arginine at a position corresponding to amino acid position 94, alanine at a position corresponding to amino acid position 104, phenylalanine at a position corresponding to amino acid position 107, phenylalanine at a position corresponding to amino acid position 107, valine at a position corresponding to amino acid position 109, isoleucine at a position corresponding to amino acid position 120, arginine at a position corresponding to amino acid position 128, isoleucine at a position corresponding to amino acid position 136, lysine at a position corresponding to amino acid position 155, methionine at a position corresponding to amino acid position 338, termination codon at a position corresponding to amino acid position 344 and cysteine at a position corresponding to amino

acid position 497. In some embodiments, the polymorphic variant also comprises at least one variant amino acid selected from the group consisting of methionine at a position corresponding to amino acid position 7, methionine at a position corresponding to amino acid position 11, serine at a position corresponding to amino acid position 34, methionine at a position corresponding to amino acid position 91, isoleucine at a position corresponding to amino acid position 136, glutamic acid at a position corresponding to amino acid position 151 and threonine at a position corresponding to amino acid position 486. A polymorphic variant of CYP2D6 is useful in studying the effect of the variation on the biological activity of CYP2D6 as well as on the binding affinity of candidate drugs to CYP2D6, or studying the enzymatic properties of such CYP2D6 variants using these candidate drugs as substrates. Herein, the term drug refers to a candidate drug or any of its metabolic derivatives.

The present invention also provides antibodies that recognize and bind to the above polymorphic CYP2D6 protein variant. Such antibodies can be utilized in a variety of diagnostic and prognostic formats and therapeutic methods.

The present invention also provides nonhuman transgenic animals comprising one or more of the CYP2D6 polymorphic genomic variants described herein and methods for producing such animals. The transgenic animals are useful for studying expression of the CYP2D6 isogenes *in vivo*, for *in vivo* screening and testing of drugs targeted against CYP2D6 protein, and for testing the efficacy of therapeutic agents and compounds for hypertension, atrial and ventricular arrhythmias Parkinson's disease and drug-induced lupus syndrome in a biological system.

The present invention also provides a computer system for storing and displaying polymorphism data determined for the CYP2D6 gene. The computer system comprises a computer processing unit; a display; and a database containing the polymorphism data. The polymorphism data includes one or more of the following: the polymorphisms, the genotypes, the haplotypes, and the haplotype pairs identified for the CYP2D6 gene in a reference population. In a preferred embodiment, the computer system is capable of producing a display showing CYP2D6 haplotypes organized according to their evolutionary relationships.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates a reference sequence for the CYP2D6 gene (Genaissance Reference No. 3578126; contiguous lines), with the start and stop positions of each region of coding sequence indicated with a bracket ([or]) and the numerical position below the sequence and the polymorphic site(s) and polymorphism(s) identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the sequence. SEQ ID NO:1 is equivalent to Figure 1, with the two alternative allelic variants of each polymorphic site indicated by the appropriate nucleotide symbol (R= G or A, Y= T or C, M= A or C, K= G or T, S= G or C, and W= A or T; WIPO standard ST.25). SEQ ID NO:151 is a modified version of SEQ ID NO:1 that shows the context sequence of each polymorphic site, PS1-PS42, in a uniform format to facilitate electronic searching.

For each polymorphic site, SEQ ID NO:151 contains a block of 60 bases of the nucleotide sequence encompassing the centrally-located polymorphic site at the 30th position, followed by 60 bases of unspecified sequence to represent that each PS is separated by genomic sequence whose composition is defined elsewhere herein.

Figure 2 illustrates a reference sequence for the CYP2D6 coding sequence (contiguous lines; SEQ ID NO:2), with the polymorphic site(s) and polymorphism(s) identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the sequence.

Figure 3 illustrates a reference sequence for the CYP2D6 protein (contiguous lines; SEQ ID NO:3), with the variant amino acid(s) caused by the polymorphism(s) of Figure 2 positioned below the polymorphic site in the sequence.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is based on the discovery of novel variants of the CYP2D6 gene. As described in more detail below, the inventors herein discovered 34 isogenes of the CYP2D6 gene by characterizing the CYP2D6 gene found in genomic DNAs isolated from an Index Repository that contains immortalized cell lines from one chimpanzee and 93 human individuals. The human individuals included a reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: Caucasian (21 individuals), African descent (20 individuals), Asian (20 individuals), or Hispanic/Latino (18 individuals). To the extent possible, the members of this reference population were organized into population subgroups by their self-identified ethnogeographic origin as shown in Table 1 below.

Table 1. Population Groups in the Index Repository

Population Group	Population Subgroup	No. of Individuals
African descent		20
	Sierra Leone	1
Asian		20
	Burma	1
	China	3
	Japan	6
	Korea	1
	Philippines	5
	Vietnam	4
Caucasian		21
	British Isles	3
	British Isles/Central	4
	British Isles/Eastern	1
	Central/Eastern	1
	Eastern	3
	Central/Mediterranean	1
	Mediterranean	2
	Scandinavian	2
Hispanic/Latino		18
	Caribbean	8
	Caribbean (Spanish Descent)	2
	Central American (Spanish Descent)	1
	Mexican American	4
	South American (Spanish Descent)	3

In addition, the Index Repository contains three unrelated indigenous American Indians (one from each of North, Central and South America), one three-generation Caucasian family (from the CEPH Utah cohort) and one two-generation African-American family.

The CYP2D6 isogenes present in the human reference population are defined by haplotypes for 42 polymorphic sites in the CYP2D6 gene, 29 of which are believed to be novel. The CYP2D6 polymorphic sites identified by the inventors are referred to as PS1-PS42 to designate the order in which they are located in the gene (see Table 2 below), with the novel polymorphic sites referred to as PS1, PS2, PS3, PS4, PS6, PS10, PS12, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS22, PS23, PS24, PS26, PS27, PS28, PS29, PS32, PS34, PS36, PS37, PS38, PS39, PS40 and PS42. Using the genotypes identified in the Index Repository for PS1-PS42 and the methodology described in the Examples below, the inventors herein also determined the pair of haplotypes for the CYP2D6 gene present in individual human members of this repository. The human genotypes and haplotypes found in the repository for the CYP2D6 gene include those shown in Tables 3 and 4, respectively. The polymorphism and haplotype data disclosed herein are useful for validating whether CYP2D6 is a suitable target for drugs to treat hypertension, atrial and ventricular arrhythmias Parkinson's disease and drug-induced lupus syndrome, screening for such drugs and reducing bias in clinical trials of such drugs.

In the context of this disclosure, the following terms shall be defined as follows unless

otherwise indicated:

Allele - A particular form of a genetic locus, distinguished from other forms by its particular nucleotide sequence.

Candidate Gene – A gene which is hypothesized to be responsible for a disease, condition, or the response to a treatment, or to be correlated with one of these.

Gene - A segment of DNA that contains all the information for the regulated biosynthesis of an RNA product, including promoters, exons, introns, and other untranslated regions that control expression.

Genotype – An unphased 5′ to 3′ sequence of nucleotide pair(s) found at one or more polymorphic sites in a locus on a pair of homologous chromosomes in an individual. As used herein, genotype includes a full-genotype and/or a sub-genotype as described below.

Full-genotype – The unphased 5′ to 3′ sequence of nucleotide pairs found at all polymorphic sites examined herein in a locus on a pair of homologous chromosomes in a single individual.

Sub-genotype – The unphased 5′ to 3′ sequence of nucleotides seen at a subset of the polymorphic sites examined herein in a locus on a pair of homologous chromosomes in a single individual.

Genotyping – A process for determining a genotype of an individual.

Haplotype – A 5′ to 3′ sequence of nucleotides found at one or more polymorphic sites in a locus on a single chromosome from a single individual. As used herein, haplotype includes a full-haplotype and/or a sub-haplotype as described below.

Full-haplotype – The 5′ to 3′ sequence of nucleotides found at all polymorphic sites examined herein in a locus on a single chromosome from a single individual.

Sub-haplotype – The 5′ to 3′ sequence of nucleotides seen at a subset of the polymorphic sites examined herein in a locus on a single chromosome from a single individual.

Haplotype pair – The two haplotypes found for a locus in a single individual.

Haplotyping – A process for determining one or more haplotypes in an individual and includes use of family pedigrees, molecular techniques and/or statistical inference.

Haplotype data - Information concerning one or more of the following for a specific gene: a listing of the haplotype pairs in each individual in a population; a listing of the different haplotypes in a population; frequency of each haplotype in that or other populations, and any known associations between one or more haplotypes and a trait.

Isoform – A particular form of a gene, mRNA, cDNA, coding sequence or the protein encoded thereby, distinguished from other forms by its particular sequence and/or structure.

Isogene – One of the isoforms (e.g., alleles) of a gene found in a population. An isogene (or allele) contains all of the polymorphisms present in the particular isoform of the gene.

Isolated – As applied to a biological molecule such as RNA, DNA, oligonucleotide, or protein, isolated means the molecule is substantially free of other biological molecules such as nucleic

acids, proteins, lipids, carbohydrates, or other material such as cellular debris and growth media. Generally, the term "isolated" is not intended to refer to a complete absence of such material or to absence of water, buffers, or salts, unless they are present in amounts that substantially interfere with the methods of the present invention.

Locus - A location on a chromosome or DNA molecule corresponding to a gene or a physical or phenotypic feature, where physical features include polymorphic sites.

Naturally-occurring – A term used to designate that the object it is applied to, e.g., naturally-occurring polynucleotide or polypeptide, can be isolated from a source in nature and which has not been intentionally modified by man.

Nucleotide pair – The nucleotides found at a polymorphic site on the two copies of a chromosome from an individual.

Phased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, phased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is known.

Polymorphic site (PS) – A position on a chromosome or DNA molecule at which at least two alternative sequences are found in a population.

Polymorphic variant (or variant)– A gene, mRNA, cDNA, polypeptide, protein or peptide whose nucleotide or amino acid sequence varies from a reference sequence due to the presence of a polymorphism in the gene.

Polymorphism – The sequence variation observed in an individual at a polymorphic site. Polymorphisms include nucleotide substitutions, insertions, deletions and microsatellites and may, but need not, result in detectable differences in gene expression or protein function.

Polymorphism data – Information concerning one or more of the following for a specific gene: location of polymorphic sites; sequence variation at those sites; frequency of polymorphisms in one or more populations; the different genotypes and/or haplotypes determined for the gene; frequency of one or more of these genotypes and/or haplotypes in one or more populations; any known association(s) between a trait and a genotype or a haplotype for the gene.

Polymorphism Database – A collection of polymorphism data arranged in a systematic or methodical way and capable of being individually accessed by electronic or other means.

Polynucleotide – A nucleic acid molecule comprised of single-stranded RNA or DNA or comprised of complementary, double-stranded DNA.

Population Group – A group of individuals sharing a common ethnogeographic origin.

Reference Population – A group of subjects or individuals who are predicted to be representative of the genetic variation found in the general population. Typically, the reference population represents the genetic variation in the population at a certainty level of at least 85%, preferably at least 90%, more preferably at least 95% and even more preferably at least 99%.

Single Nucleotide Polymorphism (SNP) – Typically, the specific pair of nucleotides

observed at a single polymorphic site. In rare cases, three or four nucleotides may be found.

Subject – A human individual whose genotypes or haplotypes or response to treatment or disease state are to be determined.

Treatment - A stimulus administered internally or externally to a subject.

Unphased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, unphased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is not known.

As discussed above, information on the identity of genotypes and haplotypes for the CYP2D6 gene of any particular individual as well as the frequency of such genotypes and haplotypes in any particular population of individuals is useful for a variety of drug discovery and development applications. Thus, the invention also provides compositions and methods for detecting the novel CYP2D6 polymorphisms, haplotypes and haplotype pairs identified herein.

The compositions comprise at least one oligonucleotide for detecting the variant nucleotide or nucleotide pair located at a novel CYP2D6 polymorphic site in one copy or two copies of the CYP2D6 gene. Such oligonucleotides are referred to herein as CYP2D6 haplotyping oligonucleotides or genotyping oligonucleotides, respectively, and collectively as CYP2D6 oligonucleotides. In one embodiment, a CYP2D6 haplotyping or genotyping oligonucleotide is a probe or primer capable of hybridizing to a target region that contains, or that is located close to, one of the novel polymorphic sites described herein.

As used herein, the term "oligonucleotide" refers to a polynucleotide molecule having less than about 100 nucleotides. A preferred oligonucleotide of the invention is 10 to 35 nucleotides long. More preferably, the oligonucleotide is between 15 and 30, and most preferably, between 20 and 25 nucleotides in length. The exact length of the oligonucleotide will depend on many factors that are routinely considered and practiced by the skilled artisan. The oligonucleotide may be comprised of any phosphorylation state of ribonucleotides, deoxyribonucleotides, and acyclic nucleotide derivatives, and other functionally equivalent derivatives. Alternatively, oligonucleotides may have a phosphate-free backbone, which may be comprised of linkages such as carboxymethyl, acetamidate, carbamate, polyamide (peptide nucleic acid (PNA)) and the like (Varma, R. in Molecular Biology and Biotechnology, A Comprehensive Desk Reference, Ed. R. Meyers, VCH Publishers, Inc. (1995), pages 617-620). Oligonucleotides of the invention may be prepared by chemical synthesis using any suitable methodology known in the art, or may be derived from a biological sample, for example, by restriction digestion. The oligonucleotides may be labeled, according to any technique known in the art, including use of radiolabels, fluorescent labels, enzymatic labels, proteins, haptens, antibodies, sequence tags and the like.

Haplotyping or genotyping oligonucleotides of the invention must be capable of specifically hybridizing to a target region of a CYP2D6 polynucleotide. Preferably, the target region is located in a CYP2D6 isogene. As used herein, specific hybridization means the oligonucleotide forms an anti-

parallel double-stranded structure with the target region under certain hybridizing conditions, while failing to form such a structure when incubated with another region in the CYP2D6 polynucleotide or with a non-CYP2D6 polynucleotide under the same hybridizing conditions. Preferably, the oligonucleotide specifically hybridizes to the target region under conventional high stringency conditions. The skilled artisan can readily design and test oligonucleotide probes and primers suitable for detecting polymorphisms in the CYP2D6 gene using the polymorphism information provided herein in conjunction with the known sequence information for the CYP2D6 gene and routine techniques.

A nucleic acid molecule such as an oligonucleotide or polynucleotide is said to be a "perfect" or "complete" complement of another nucleic acid molecule if every nucleotide of one of the molecules is complementary to the nucleotide at the corresponding position of the other molecule. A nucleic acid molecule is "substantially complementary" to another molecule if it hybridizes to that molecule with sufficient stability to remain in a duplex form under conventional low-stringency conditions. Conventional hybridization conditions are described, for example, by Sambrook J. et al., in Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989) and by Haymes, B.D. et al. in Nucleic Acid Hybridization, A Practical Approach, IRL Press, Washington, D.C. (1985). While perfectly complementary oligonucleotides are preferred for detecting polymorphisms, departures from complete complementarity are contemplated where such departures do not prevent the molecule from specifically hybridizing to the target region. For example, an oligonucleotide primer may have a non-complementary fragment at its 5′ end, with the remainder of the primer being complementary to the target region. Alternatively, non-complementary nucleotides may be interspersed into the probe or primer as long as the resulting probe or primer is still capable of specifically hybridizing to the target region.

Preferred haplotyping or genotyping oligonucleotides of the invention are allele-specific oligonucleotides. As used herein, the term allele-specific oligonucleotide (ASO) means an oligonucleotide that is able, under sufficiently stringent conditions, to hybridize specifically to one allele of a gene, or other locus, at a target region containing a polymorphic site while not hybridizing to the corresponding region in another allele(s). As understood by the skilled artisan, allele-specificity will depend upon a variety of readily optimized stringency conditions, including salt and formamide concentrations, as well as temperatures for both the hybridization and washing steps. Examples of hybridization and washing conditions typically used for ASO probes are found in Kogan et al., "Genetic Prediction of Hemophilia A" in PCR Protocols, A Guide to Methods and Applications, Academic Press, 1990 and Ruaño et al., 87 *Proc. Natl. Acad. Sci. USA* 6296-6300, 1990. Typically, an ASO will be perfectly complementary to one allele while containing a single mismatch for another allele.

Allele-specific oligonucleotides of the invention include ASO probes and ASO primers. ASO probes which usually provide good discrimination between different alleles are those in which a

central position of the oligonucleotide probe aligns with the polymorphic site in the target region (e.g., approximately the 7th or 8th position in a 15mer, the 8th or 9th position in a 16mer, and the 10th or 11th position in a 20mer). An ASO primer of the invention has a 3′ terminal nucleotide, or preferably a 3′ penultimate nucleotide, that is complementary to only one nucleotide of a particular SNP, thereby acting as a primer for polymerase-mediated extension only if the allele containing that nucleotide is present. ASO probes and primers hybridizing to either the coding or noncoding strand are contemplated by the invention. ASO probes and primers listed below use the appropriate nucleotide symbol (R= G or A, Y= T or C, M= A or C, K= G or T, S= G or C, and W= A or T; WIPO standard ST.25) at the position of the polymorphic site to represent that the ASO contains either of the two alternative allelic variants observed at that polymorphic site.

A preferred ASO probe for detecting CYP2D6 gene polymorphisms comprises a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

GATGGCCRGGTCCAC (SEQ ID NO:4) and its complement,
TGCAGGCYTCAGGAG (SEQ ID NO:5) and its complement,
GAAGCAGSGGCAAGA (SEQ ID NO:6) and its complement,
GGGCAAGRACCTCTG (SEQ ID NO:7) and its complement,
GGTGTGCYGAGAGTG (SEQ ID NO:8) and its complement,
GGACTAGSACCTGTA (SEQ ID NO:9) and its complement,
GCCGTGCRCGAGGCG (SEQ ID NO:10) and its complement,
GTGACCCRCGGCGAG (SEQ ID NO:11) and its complement,
AGGACACSGCCGACC (SEQ ID NO:12) and its complement,
CCGCCTGYGCCCATC (SEQ ID NO:13) and its complement,
GCCCATCWCCCAGAT (SEQ ID NO:14) and its complement,
CCCATCAYCCAGATC (SEQ ID NO:15) and its complement,
CACCCAGRTCCTGGG (SEQ ID NO:16) and its complement,
TCCTGGGYTTCGGGC (SEQ ID NO:17) and its complement,
GGCAAGCRGCGGTGG (SEQ ID NO:18) and its complement,
GCAGCGGKGGGGACA (SEQ ID NO:19) and its complement,
CGTAGTCYGAGCTGG (SEQ ID NO:20) and its complement,
GGACAGCSGGCCAAG (SEQ ID NO:21) and its complement,
AGGGGTGWTCCTGGC (SEQ ID NO:22) and its complement,
GCCCGCGYGGCGCGA (SEQ ID NO:23) and its complement,
CTTCTCCRTGTCCAC (SEQ ID NO:24) and its complement,
GGTGACCRAGGAGGC (SEQ ID NO:25) and its complement,
AGAGAGGRTGGAGGC (SEQ ID NO:26) and its complement,
CGACGACRTGATAGG (SEQ ID NO:27) and its complement,
GGTGCGGYGACCAGA (SEQ ID NO:28) and its complement,
TGATTCAYGAGGTGC (SEQ ID NO:29) and its complement,
CACCAGCMCCTGGTG (SEQ ID NO:30) and its complement,
CTAGGAAMCCTGGCC (SEQ ID NO:31) and its complement, and
TGTGCCCYGCTAGAA (SEQ ID NO:32) and its complement.

A preferred ASO primer for detecting CYP2D6 gene polymorphisms comprises a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

GAGGTGGATGGCCRG (SEQ ID NO:33); GTTTCAGTGGACCYG (SEQ ID NO:34);
AGGCTTTGCAGGCYT (SEQ ID NO:35); TCCAAGCTCCTGARG (SEQ ID NO:36);
GAGCAGGAAGCAGSG (SEQ ID NO:37); AGAGGTTCTTGCCSC (SEQ ID NO:38);
AAGCAGGGGCAAGRA (SEQ ID NO:39); CTGCTCCAGAGGTYC (SEQ ID NO:40);
GCGCTCGGTGTGCYG (SEQ ID NO:41); GCAGGACACTCTCRG (SEQ ID NO:42);
CAAATAGGACTAGSA (SEQ ID NO:43); CCAGACTACAGGTSC (SEQ ID NO:44);
CTGGCGGCCGTGCRC (SEQ ID NO:45); CACCAGCGCCTCGYG (SEQ ID NO:46);
GCGCTGGTGACCCRC (SEQ ID NO:47); GGTGTCCTCGCCGYG (SEQ ID NO:48);
ACGGCGAGGACACSG (SEQ ID NO:49); GCGGGCGGTCGGCSG (SEQ ID NO:50);
GACCGCCCGCCTGYG (SEQ ID NO:51); CTGGGTGATGGGCRC (SEQ ID NO:52);
GCCTGTGCCCATCWC (SEQ ID NO:53); CCCAGGATCTGGGWG (SEQ ID NO:54);
CCTGTGCCCATCAYC (SEQ ID NO:55); ACCCAGGATCTGGRT (SEQ ID NO:56);
GCCCATCACCCAGRT (SEQ ID NO:57); CCGAAACCCAGGAYC (SEQ ID NO:58);
CCCAGATCCTGGGYT (SEQ ID NO:59); AACGCGGCCCGAARC (SEQ ID NO:60);
TCCCAAGGCAAGCRG (SEQ ID NO:61); CTGTCCCCACCGCYG (SEQ ID NO:62);
AGGCAAGCAGCGGKG (SEQ ID NO:63); TGTCTCTGTCCCCMC (SEQ ID NO:64);
GATGACCGTAGTCYG (SEQ ID NO:65); CTCTGCCCAGCTCRG (SEQ ID NO:66);
GAGTGGGGACAGCSG (SEQ ID NO:67); TGGTTTCTTGGCCSG (SEQ ID NO:68);
ACCCCCAGGGGTGWT (SEQ ID NO:69); TAGCGCGCCAGGAWC (SEQ ID NO:70);
CTATGGGCCCGCGYG (SEQ ID NO:71); CTCTGCTCGCGCCRC (SEQ ID NO:72);
GAGGCGCTTCTCCRT (SEQ ID NO:73); CGCAAGGTGGACAYG (SEQ ID NO:74);
GCAGTGGGTGACCRA (SEQ ID NO:75); CAGGCGGCCTCCTYG (SEQ ID NO:76);
GCAAGGAGAGAGGRT (SEQ ID NO:77); GTGCCAGCCTCCAYC (SEQ ID NO:78);
GGAGATCGACGACRT (SEQ ID NO:79); ACCTGCCCTATCAYG (SEQ ID NO:80);
AGGGCAGGTGCGGYG (SEQ ID NO:81); CCCATCTCTGGTCRC (SEQ ID NO:82);
CTGCCGTGATTCAYG (SEQ ID NO:83); AGCGCTGCACCTCRT (SEQ ID NO:84);
GCTCAGCACCAGCMC (SEQ ID NO:85); GGCTATCACCAGGKG (SEQ ID NO:86);
CCCACTCTAGGAAMC (SEQ ID NO:87); CTAGGTGGCCAGGKT (SEQ ID NO:88);
TTGTGCTGTGCCCYG (SEQ ID NO:89); and ACCCCATTCTAGCRG (SEQ ID NO:90).

Other oligonucleotides of the invention hybridize to a target region located one to several nucleotides downstream of one of the novel polymorphic sites identified herein. Such oligonucleotides are useful in polymerase-mediated primer extension methods for detecting one of the novel polymorphisms described herein and therefore such oligonucleotides are referred to herein as "primer-extension oligonucleotides". In a preferred embodiment, the 3′-terminus of a primer-extension oligonucleotide is a deoxynucleotide complementary to the nucleotide located immediately adjacent to the polymorphic site.

A particularly preferred oligonucleotide primer for detecting CYP2D6 gene polymorphisms by primer extension terminates in a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

GTGGATGGCC (SEQ ID NO:91); TCAGTGGACC (SEQ ID NO:92);
CTTTGCAGGC (SEQ ID NO:93); AAGCTCCTGA (SEQ ID NO:94);
CAGGAAGCAG (SEQ ID NO:95); GGTTCTTGCC (SEQ ID NO:96);
CAGGGGCAAG (SEQ ID NO:97); CTCCAGAGGT (SEQ ID NO:98);
CTCGGTGTGC (SEQ ID NO:99); GGACACTCTC (SEQ ID NO:100);
ATAGGACTAG (SEQ ID NO:101); GACTACAGGT (SEQ ID NO:102);
GCGGCCGTGC (SEQ ID NO:103); CAGCGCCTCG (SEQ ID NO:104);
CTGGTGACCC (SEQ ID NO:105); GTCCTCGCCG (SEQ ID NO:106);
GCGAGGACAC (SEQ ID NO:107); GGCGGTCGGC (SEQ ID NO:108);

CGCCCGCCTG (SEQ ID NO:109); GGTGATGGGC (SEQ ID NO:110);
TGTGCCCATC (SEQ ID NO:111); AGGATCTGGG (SEQ ID NO:112);
GTGCCCATCA (SEQ ID NO:113); CAGGATCTGG (SEQ ID NO:114);
CATCACCCAG (SEQ ID NO:115); AAACCCAGGA (SEQ ID NO:116);
AGATCCTGGG (SEQ ID NO:117); GCGGCCCGAA (SEQ ID NO:118);
CAAGGCAAGC (SEQ ID NO:119); TCCCCACCGC (SEQ ID NO:120);
CAAGCAGCGG (SEQ ID NO:121); CTCTGTCCCC (SEQ ID NO:122);
GACCGTAGTC (SEQ ID NO:123); TGCCCAGCTC (SEQ ID NO:124);
TGGGGACAGC (SEQ ID NO:125); TTTCTTGGCC (SEQ ID NO:126);
CCCAGGGGTG (SEQ ID NO:127); CGCGCCAGGA (SEQ ID NO:128);
TGGGCCCGCG (SEQ ID NO:129); TGCTCGCGCC (SEQ ID NO:130);
GCGCTTCTCC (SEQ ID NO:131); AAGGTGGACA (SEQ ID NO:132);
GTGGGTGACC (SEQ ID NO:133); GCGGCCTCCT (SEQ ID NO:134);
AGGAGAGAGG (SEQ ID NO:135); CCAGCCTCCA (SEQ ID NO:136);
GATCGACGAC (SEQ ID NO:137); TGCCCTATCA (SEQ ID NO:138);
GCAGGTGCGG (SEQ ID NO:139); ATCTCTGGTC (SEQ ID NO:140);
CCGTGATTCA (SEQ ID NO:141); GCTGCACCTC (SEQ ID NO:142);
CAGCACCAGC (SEQ ID NO:143); TATCACCAGG (SEQ ID NO:144);
ACTCTAGGAA (SEQ ID NO:145); GGTGGCCAGG (SEQ ID NO:146);
TGCTGTGCCC (SEQ ID NO:147); and CCATTCTAGC (SEQ ID NO:148).

In some embodiments, a composition contains two or more differently labeled CYP2D6 oligonucleotides for simultaneously probing the identity of nucleotides or nucleotide pairs at two or more polymorphic sites. It is also contemplated that primer compositions may contain two or more sets of allele-specific primer pairs to allow simultaneous targeting and amplification of two or more regions containing a polymorphic site.

CYP2D6 oligonucleotides of the invention may also be immobilized on or synthesized on a solid surface such as a microchip, bead, or glass slide (see, e.g., WO 98/20020 and WO 98/20019). Such immobilized oligonucleotides may be used in a variety of polymorphism detection assays, including but not limited to probe hybridization and polymerase extension assays. Immobilized CYP2D6 oligonucleotides of the invention may comprise an ordered array of oligonucleotides designed to rapidly screen a DNA sample for polymorphisms in multiple genes at the same time.

In another embodiment, the invention provides a kit comprising at least two CYP2D6 oligonucleotides packaged in separate containers. The kit may also contain other components such as hybridization buffer (where the oligonucleotides are to be used as a probe) packaged in a separate container. Alternatively, where the oligonucleotides are to be used to amplify a target region, the kit may contain, packaged in separate containers, a polymerase and a reaction buffer optimized for primer extension mediated by the polymerase, such as PCR.

The above described oligonucleotide compositions and kits are useful in methods for genotyping and/or haplotyping the CYP2D6 gene in an individual. As used herein, the terms "CYP2D6 genotype" and "CYP2D6 haplotype" mean the genotype or haplotype contains the nucleotide pair or nucleotide, respectively, that is present at one or more of the novel polymorphic sites described herein and may optionally also include the nucleotide pair or nucleotide present at one or more additional polymorphic sites in the CYP2D6 gene. The additional polymorphic sites may be

currently known polymorphic sites or sites that are subsequently discovered.

One embodiment of a genotyping method of the invention involves isolating from the individual a nucleic acid sample comprising the two copies of the CYP2D6 gene, mRNA transcripts thereof or cDNA copies thereof, or a fragment of any of the foregoing, that are present in the individual, and determining the identity of the nucleotide pair at one or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS6, PS10, PS12, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS22, PS23, PS24, PS26, PS27, PS28, PS29, PS32, PS34, PS36, PS37, PS38, PS39, PS40 and PS42 in the two copies to assign a CYP2D6 genotype to the individual. As will be readily understood by the skilled artisan, the two "copies" of a gene, mRNA or cDNA (or fragment of such CYP2D6 molecules) in an individual may be the same allele or may be different alleles. In a preferred embodiment of the method for assigning a CYP2D6 genotype, the identity of the nucleotide pair at one or more of the polymorphic sites selected from the group consisting of PS5, PS7, PS8, PS9, PS11, PS13, PS21, PS25, PS30, PS31, PS33, PS35 and PS41 is also determined. In another embodiment, a genotyping method of the invention comprises determining the identity of the nucleotide pair at each of PS1-PS42.

Typically, the nucleic acid sample is isolated from a biological sample taken from the individual, such as a blood sample or tissue sample. Suitable tissue samples include whole blood, semen, saliva, tears, urine, fecal material, sweat, buccal, skin and hair. The nucleic acid sample may be comprised of genomic DNA, mRNA, or cDNA and, in the latter two cases, the biological sample must be obtained from a tissue in which the CYP2D6 gene is expressed. Furthermore it will be understood by the skilled artisan that mRNA or cDNA preparations would not be used to detect polymorphisms located in introns or in 5′ and 3′ untranslated regions if not present in the mRNA or cDNA. If a CYP2D6 gene fragment is isolated, it must contain the polymorphic site(s) to be genotyped.

One embodiment of a haplotyping method of the invention comprises isolating from the individual a nucleic acid sample containing only one of the two copies of the CYP2D6 gene, mRNA or cDNA, or a fragment of such CYP2D6 molecules, that is present in the individual and determining in that copy the identity of the nucleotide at one or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS6, PS10, PS12, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS22, PS23, PS24, PS26, PS27, PS28, PS29, PS32, PS34, PS36, PS37, PS38, PS39, PS40 and PS42 in that copy to assign a CYP2D6 haplotype to the individual.

The nucleic acid used in the above haplotyping methods of the invention may be isolated using any method capable of separating the two copies of the CYP2D6 gene or fragment such as one of the methods described above for preparing CYP2D6 isogenes, with targeted *in vivo* cloning being the preferred approach. As will be readily appreciated by those skilled in the art, any individual clone will typically only provide haplotype information on one of the two CYP2D6 gene copies present in an individual. If haplotype information is desired for the individual's other copy, additional CYP2D6

clones will usually need to be examined. Typically, at least five clones should be examined to have more than a 90% probability of haplotyping both copies of the CYP2D6 gene in an individual. In some cases, however, once the haplotype for one CYP2D6 allele is directly determined, the haplotype for the other allele may be inferred if the individual has a known genotype for the polymorphic sites of interest or if the haplotype frequency or haplotype pair frequency for the individual's population group is known. In some embodiments, the CYP2D6 haplotype is assigned to the individual by also identifying the nucleotide at one or more polymorphic sites selected from the group consisting of PS5, PS7, PS8, PS9, PS11, PS13, PS21, PS25, PS30, PS31, PS33, PS35 and PS41. In a particularly preferred embodiment, the nucleotide at each of PS1-PS42 is identified.

In another embodiment, the haplotyping method comprises determining whether an individual has one or more of the CYP2D6 haplotypes shown in Table 4. This can be accomplished by identifying, for one or both copies of the individual's CYP2D6 gene, the phased sequence of nucleotides present at each of PS1-PS42. This identifying step does not necessarily require that each of PS1-PS42 be directly examined. Typically only a subset of PS1-PS42 will need to be directly examined to assign to an individual one or more of the haplotypes shown in Table 4. This is because at least one polymorphic site in a gene is frequently in strong linkage disequilibrium with one or more other polymorphic sites in that gene (Drysdale, CM et al. 2000 *PNAS* **97**:10483-10488; Rieder MJ et al. 1999 *Nature Genetics* **22**:59-62). Two sites are said to be in linkage disequilibrium if the presence of a particular variant at one site enhances the predictability of another variant at the second site (Stephens, JC 1999, *Mol. Diag.* **4**:309-317). Techniques for determining whether any two polymorphic sites are in linkage disequilibrium are well-known in the art (Weir B.S. 1996 *Genetic Data Analysis II*, Sinauer Associates, Inc. Publishers, Sunderland, MA).

In another embodiment of a haplotyping method of the invention, a CYP2D6 haplotype pair is determined for an individual by identifying the phased sequence of nucleotides at one or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS6, PS10, PS12, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS22, PS23, PS24, PS26, PS27, PS28, PS29, PS32, PS34, PS36, PS37, PS38, PS39, PS40 and PS42 in each copy of the CYP2D6 gene that is present in the individual. In a particularly preferred embodiment, the haplotyping method comprises identifying the phased sequence of nucleotides at each of PS1-PS42 in each copy of the CYP2D6 gene.

When haplotyping both copies of the gene, the identifying step is preferably performed with each copy of the gene being placed in separate containers. However, it is also envisioned that if the two copies are labeled with different tags, or are otherwise separately distinguishable or identifiable, it could be possible in some cases to perform the method in the same container. For example, if first and second copies of the gene are labeled with different first and second fluorescent dyes, respectively, and an allele-specific oligonucleotide labeled with yet a third different fluorescent dye is used to assay the polymorphic site(s), then detecting a combination of the first and third dyes would identify the polymorphism in the first gene copy while detecting a combination of the second and third dyes would

identify the polymorphism in the second gene copy.

In both the genotyping and haplotyping methods, the identity of a nucleotide (or nucleotide pair) at a polymorphic site(s) may be determined by amplifying a target region(s) containing the polymorphic site(s) directly from one or both copies of the CYP2D6 gene, or a fragment thereof, and the sequence of the amplified region(s) determined by conventional methods. It will be readily appreciated by the skilled artisan that only one nucleotide will be detected at a polymorphic site in individuals who are homozygous at that site, while two different nucleotides will be detected if the individual is heterozygous for that site. The polymorphism may be identified directly, known as positive-type identification, or by inference, referred to as negative-type identification. For example, where a SNP is known to be guanine and cytosine in a reference population, a site may be positively determined to be either guanine or cytosine for an individual homozygous at that site, or both guanine and cytosine, if the individual is heterozygous at that site. Alternatively, the site may be negatively determined to be not guanine (and thus cytosine/cytosine) or not cytosine (and thus guanine/guanine).

The target region(s) may be amplified using any oligonucleotide-directed amplification method, including but not limited to polymerase chain reaction (PCR) (U.S. Patent No. 4,965,188), ligase chain reaction (LCR) (Barany et al., *Proc. Natl. Acad. Sci. USA* 88:189-193, 1991; WO90/01069), and oligonucleotide ligation assay (OLA) (Landegren et al., *Science* 241:1077-1080, 1988). Other known nucleic acid amplification procedures may be used to amplify the target region including transcription-based amplification systems (U.S. Patent No. 5,130,238; EP 329,822; U.S. Patent No. 5,169,766, WO89/06700) and isothermal methods (Walker et al., *Proc. Natl. Acad. Sci. USA* **89**:392-396, 1992).

A polymorphism in the target region may also be assayed before or after amplification using one of several hybridization-based methods known in the art. Typically, allele-specific oligonucleotides are utilized in performing such methods. The allele-specific oligonucleotides may be used as differently labeled probe pairs, with one member of the pair showing a perfect match to one variant of a target sequence and the other member showing a perfect match to a different variant. In some embodiments, more than one polymorphic site may be detected at once using a set of allele-specific oligonucleotides or oligonucleotide pairs. Preferably, the members of the set have melting temperatures within 5°C, and more preferably within 2°C, of each other when hybridizing to each of the polymorphic sites being detected.

Hybridization of an allele-specific oligonucleotide to a target polynucleotide may be performed with both entities in solution, or such hybridization may be performed when either the oligonucleotide or the target polynucleotide is covalently or noncovalently affixed to a solid support. Attachment may be mediated, for example, by antibody-antigen interactions, poly-L-Lys, streptavidin or avidin-biotin, salt bridges, hydrophobic interactions, chemical linkages, UV cross-linking baking, etc. Allele-specific oligonucleotides may be synthesized directly on the solid support or attached to the solid support subsequent to synthesis. Solid-supports suitable for use in detection methods of the

invention include substrates made of silicon, glass, plastic, paper and the like, which may be formed, for example, into wells (as in 96-well plates), slides, sheets, membranes, fibers, chips, dishes, and beads. The solid support may be treated, coated or derivatized to facilitate the immobilization of the allele-specific oligonucleotide or target nucleic acid.

The genotype or haplotype for the CYP2D6 gene of an individual may also be determined by hybridization of a nucleic acid sample containing one or both copies of the gene, mRNA, cDNA or fragment(s) thereof, to nucleic acid arrays and subarrays such as described in WO 95/11995. The arrays would contain a battery of allele-specific oligonucleotides representing each of the polymorphic sites to be included in the genotype or haplotype.

The identity of polymorphisms may also be determined using a mismatch detection technique, including but not limited to the RNase protection method using riboprobes (Winter et al., *Proc. Natl. Acad. Sci. USA* 82:7575, 1985; Meyers et al., *Science* 230:1242, 1985) and proteins which recognize nucleotide mismatches, such as the E. coli mutS protein (Modrich, P. *Ann. Rev. Genet.* 25:229-253, 1991). Alternatively, variant alleles can be identified by single strand conformation polymorphism (SSCP) analysis (Orita et al., *Genomics* 5:874-879, 1989; Humphries et al., in Molecular Diagnosis of Genetic Diseases, R. Elles, ed., pp. 321-340, 1996) or denaturing gradient gel electrophoresis (DGGE) (Wartell et al., *Nucl. Acids Res.* 18:2699-2706, 1990; Sheffield et al., *Proc. Natl. Acad. Sci. USA* 86:232-236, 1989).

A polymerase-mediated primer extension method may also be used to identify the polymorphism(s). Several such methods have been described in the patent and scientific literature and include the "Genetic Bit Analysis" method (WO92/15712) and the ligase/polymerase mediated genetic bit analysis (U.S. Patent 5,679,524. Related methods are disclosed in WO91/02087, WO90/09455, WO95/17676, U.S. Patent Nos. 5,302,509, and 5,945,283. Extended primers containing a polymorphism may be detected by mass spectrometry as described in U.S. Patent No. 5,605,798. Another primer extension method is allele-specific PCR (Ruaño et al., *Nucl. Acids Res.* 17:8392, 1989; Ruaño et al., *Nucl. Acids Res.* 19, 6877-6882, 1991; WO 93/22456; Turki et al., *J. Clin. Invest.* 95:1635-1641, 1995). In addition, multiple polymorphic sites may be investigated by simultaneously amplifying multiple regions of the nucleic acid using sets of allele-specific primers as described in Wallace et al. (WO89/10414).

In addition, the identity of the allele(s) present at any of the novel polymorphic sites described herein may be indirectly determined by haplotyping or genotyping another polymorphic site that is in linkage disequilibrium with the polymorphic site that is of interest. Polymorphic sites in linkage disequilibrium with the presently disclosed polymorphic sites may be located in regions of the gene or in other genomic regions not examined herein. Detection of the allele(s) present at a polymorphic site in linkage disequilibrium with the novel polymorphic sites described herein may be performed by, but is not limited to, any of the above-mentioned methods for detecting the identity of the allele at a polymorphic site.

In another aspect of the invention, an individual's CYP2D6 haplotype pair is predicted from its CYP2D6 genotype using information on haplotype pairs known to exist in a reference population. In its broadest embodiment, the haplotyping prediction method comprises identifying a CYP2D6 genotype for the individual at two or more CYP2D6 polymorphic sites described herein, accessing data containing CYP2D6 haplotype pairs identified in a reference population, and assigning a haplotype pair to the individual that is consistent with the genotype data. In one embodiment, the reference haplotype pairs include the CYP2D6 haplotype pairs shown in Table 3. The CYP2D6 haplotype pair can be assigned by comparing the individual's genotype with the genotypes corresponding to the haplotype pairs known to exist in the general population or in a specific population group, and determining which haplotype pair is consistent with the genotype of the individual. In some embodiments, the comparing step may be performed by visual inspection (for example, by consulting Table 3). When the genotype of the individual is consistent with more than one haplotype pair, frequency data (such as that presented in Table 6) may be used to determine which of these haplotype pairs is most likely to be present in the individual. This determination may also be performed in some embodiments by visual inspection, for example by consulting Table 6. If a particular CYP2D6 haplotype pair consistent with the genotype of the individual is more frequent in the reference population than others consistent with the genotype, then that haplotype pair with the highest frequency is the most likely to be present in the individual. In other embodiments, the comparison may be made by a computer-implemented algorithm with the genotype of the individual and the reference haplotype data stored in computer-readable formats. For example, as described in PCT/US01/12831, filed April 18, 2001, one computer-implemented algorithm to perform this comparison entails enumerating all possible haplotype pairs which are consistent with the genotype, accessing data containing CYP2D6 haplotype pairs frequency data determined in a reference population to determine a probability that the individual has a possible haplotype pair, and analyzing the determined probabilities to assign a haplotype pair to the individual.

Generally, the reference population should be composed of randomly-selected individuals representing the major ethnogeographic groups of the world. A preferred reference population for use in the methods of the present invention comprises an approximately equal number of individuals from Caucasian, African-descent, Asian and Hispanic-Latino population groups with the minimum number of each group being chosen based on how rare a haplotype one wants to be guaranteed to see. For example, if one wants to have a q% chance of not missing a haplotype that exists in the population at a p% frequency of occurring in the reference population, the number of individuals (n) who must be sampled is given by $2n=\log(1-q)/\log(1-p)$ where p and q are expressed as fractions. A preferred reference population allows the detection of any haplotype whose frequency is at least 10% with about 99% certainty and comprises about 20 unrelated individuals from each of the four population groups named above. A particularly preferred reference population includes a 3-generation family representing one or more of the four population groups to serve as controls for checking quality of

haplotyping procedures.

In a preferred embodiment, the haplotype frequency data for each ethnogeographic group is examined to determine whether it is consistent with Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium (D.L. Hartl et al., Principles of Population Genomics, Sinauer Associates (Sunderland, MA), 3rd Ed., 1997) postulates that the frequency of finding the haplotype pair H_1 / H_2 is equal to $p_{H-W}(H_1 / H_2) = 2p(H_1)p(H_2)$ if $H_1 \neq H_2$ and $p_{H-W}(H_1 / H_2) = p(H_1)p(H_2)$ if $H_1 = H_2$. A statistically significant difference between the observed and expected haplotype frequencies could be due to one or more factors including significant inbreeding in the population group, strong selective pressure on the gene, sampling bias, and/or errors in the genotyping process. If large deviations from Hardy-Weinberg equilibrium are observed in an ethnogeographic group, the number of individuals in that group can be increased to see if the deviation is due to a sampling bias. If a larger sample size does not reduce the difference between observed and expected haplotype pair frequencies, then one may wish to consider haplotyping the individual using a direct haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), single molecule dilution, or allele-specific long-range PCR (Michalotos-Beloin et al., *Nucleic Acids Res.* 24:4841-4843, 1996).

In one embodiment of this method for predicting a CYP2D6 haplotype pair for an individual, the assigning step involves performing the following analysis. First, each of the possible haplotype pairs is compared to the haplotype pairs in the reference population. Generally, only one of the haplotype pairs in the reference population matches a possible haplotype pair and that pair is assigned to the individual. Occasionally, only one haplotype represented in the reference haplotype pairs is consistent with a possible haplotype pair for an individual, and in such cases the individual is assigned a haplotype pair containing this known haplotype and a new haplotype derived by subtracting the known haplotype from the possible haplotype pair. Alternatively, the haplotype pair in an individual may be predicted from the individual's genotype for that gene using reported methods (e.g., Clark et al. 1990 *Mol Bio Evol* 7:111-22; copending PCT/US01/12831 filed April 18, 2001) or through a commercial haplotyping service such as offered by Genaissance Pharmaceuticals, Inc. (New Haven, CT). In rare cases, either no haplotypes in the reference population are consistent with the possible haplotype pairs, or alternatively, multiple reference haplotype pairs are consistent with the possible haplotype pairs. In such cases, the individual is preferably haplotyped using a direct molecular haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., *supra*).

The invention also provides a method for determining the frequency of a CYP2D6 genotype, haplotype, or haplotype pair in a population. The method comprises, for each member of the population, determining the genotype or the haplotype pair for the novel CYP2D6 polymorphic sites described herein, and calculating the frequency any particular genotype, haplotype, or haplotype pair is found in the population. The population may be e.g., a reference population, a family population, a

same gender population, a population group, or a trait population (e.g., a group of individuals exhibiting a trait of interest such as a medical condition or response to a therapeutic treatment).

In another aspect of the invention, frequency data for CYP2D6 genotypes, haplotypes, and/or haplotype pairs are determined in a reference population and used in a method for identifying an association between a trait and a CYP2D6 genotype, haplotype, or haplotype pair. The trait may be any detectable phenotype, including but not limited to susceptibility to a disease or response to a treatment. In one embodiment, the method involves obtaining data on the frequency of the genotype(s), haplotype(s), or haplotype pair(s) of interest in a reference population as well as in a population exhibiting the trait. Frequency data for one or both of the reference and trait populations may be obtained by genotyping or haplotyping each individual in the populations using one or more of the methods described above. The haplotypes for the trait population may be determined directly or, alternatively, by a predictive genotype to haplotype approach as described above. In another embodiment, the frequency data for the reference and/or trait populations is obtained by accessing previously determined frequency data, which may be in written or electronic form. For example, the frequency data may be present in a database that is accessible by a computer. Once the frequency data is obtained, the frequencies of the genotype(s), haplotype(s), or haplotype pair(s) of interest in the reference and trait populations are compared. In a preferred embodiment, the frequencies of all genotypes, haplotypes, and/or haplotype pairs observed in the populations are compared. If a particular CYP2D6 genotype, haplotype, or haplotype pair is more frequent in the trait population than in the reference population at a statistically significant amount, then the trait is predicted to be associated with that CYP2D6 genotype, haplotype or haplotype pair. Preferably, the CYP2D6 genotype, haplotype, or haplotype pair being compared in the trait and reference populations is selected from the full-genotypes and full-haplotypes shown in Tables 3 and 4, or from sub-genotypes and sub-haplotypes derived from these genotypes and haplotypes. Sub-genotypes useful in the invention preferably do not include sub-genotypes solely for any one of PS5, PS7, PS8, PS9, PS11, PS13, PS21, PS25, PS30, PS31, PS33, PS35 and PS41 or for any combination thereof.

In a preferred embodiment of the method, the trait of interest is a clinical response exhibited by a patient to some therapeutic treatment, for example, response to a drug targeting CYP2D6 or response to a therapeutic treatment for a medical condition. As used herein, "medical condition" includes but is not limited to any condition or disease manifested as one or more physical and/or psychological symptoms for which treatment is desirable, and includes previously and newly identified diseases and other disorders. As used herein the term "clinical response" means any or all of the following: a quantitative measure of the response, no response, and/or adverse response (i.e., side effects).

In order to deduce a correlation between clinical response to a treatment and a CYP2D6 genotype, haplotype, or haplotype pair, it is necessary to obtain data on the clinical responses exhibited by a population of individuals who received the treatment, hereinafter the "clinical

population". This clinical data may be obtained by analyzing the results of a clinical trial that has already been run and/or the clinical data may be obtained by designing and carrying out one or more new clinical trials. As used herein, the term "clinical trial" means any research study designed to collect clinical data on responses to a particular treatment, and includes but is not limited to phase I, phase II and phase III clinical trials. Standard methods are used to define the patient population and to enroll subjects.

It is preferred that the individuals included in the clinical population have been graded for the existence of the medical condition of interest. This is important in cases where the symptom(s) being presented by the patients can be caused by more than one underlying condition, and where treatment of the underlying conditions are not the same. An example of this would be where patients experience breathing difficulties that are due to either asthma or respiratory infections. If both sets were treated with an asthma medication, there would be a spurious group of apparent non-responders that did not actually have asthma. These people would affect the ability to detect any correlation between haplotype and treatment outcome. This grading of potential patients could employ a standard physical exam or one or more lab tests. Alternatively, grading of patients could use haplotyping for situations where there is a strong correlation between haplotype pair and disease susceptibility or severity.

The therapeutic treatment of interest is administered to each individual in the trial population and each individual's response to the treatment is measured using one or more predetermined criteria. It is contemplated that in many cases, the trial population will exhibit a range of responses and that the investigator will choose the number of responder groups (e.g., low, medium, high) made up by the various responses. In addition, the CYP2D6 gene for each individual in the trial population is genotyped and/or haplotyped, which may be done before or after administering the treatment.

After both the clinical and polymorphism data have been obtained, correlations between individual response and CYP2D6 genotype or haplotype content are created. Correlations may be produced in several ways. In one method, individuals are grouped by their CYP2D6 genotype or haplotype (or haplotype pair) (also referred to as a polymorphism group), and then the averages and standard deviations of clinical responses exhibited by the members of each polymorphism group are calculated.

These results are then analyzed to determine if any observed variation in clinical response between polymorphism groups is statistically significant. Statistical analysis methods which may be used are described in L.D. Fisher and G. vanBelle,."Biostatistics: A Methodology for the Health Sciences", Wiley-Interscience (New York) 1993. This analysis may also include a regression calculation of which polymorphic sites in the CYP2D6 gene give the most significant contribution to the differences in phenotype. One regression model useful in the invention is described in WO 01/01218, entitled "Methods for Obtaining and Using Haplotype Data".

A second method for finding correlations between CYP2D6 haplotype content and clinical responses uses predictive models based on error-minimizing optimization algorithms. One of many

possible optimization algorithms is a genetic algorithm (R. Judson, "Genetic Algorithms and Their Uses in Chemistry" in Reviews in Computational Chemistry, Vol. 10, pp. 1-73, K. B. Lipkowitz and D. B. Boyd, eds. (VCH Publishers, New York, 1997). Simulated annealing (Press et al., "Numerical Recipes in C: The Art of Scientific Computing", Cambridge University Press (Cambridge) 1992, Ch. 10), neural networks (E. Rich and K. Knight, "Artificial Intelligence", 2nd Edition (McGraw-Hill, New York, 1991, Ch. 18), standard gradient descent methods (Press et al., *supra,* Ch. 10), or other global or local optimization approaches (see discussion in Judson, *supra*) could also be used. Preferably, the correlation is found using a genetic algorithm approach as described in WO 01/01218.

Correlations may also be analyzed using analysis of variation (ANOVA) techniques to determine how much of the variation in the clinical data is explained by different subsets of the polymorphic sites in the CYP2D6 gene. As described in WO 01/01218, ANOVA is used to test hypotheses about whether a response variable is caused by or correlated with one or more traits or variables that can be measured (Fisher and vanBelle, *supra*, Ch. 10).

From the analyses described above, a mathematical model may be readily constructed by the skilled artisan that predicts clinical response as a function of CYP2D6 genotype or haplotype content. Preferably, the model is validated in one or more follow-up clinical trials designed to test the model.

The identification of an association between a clinical response and a genotype or haplotype (or haplotype pair) for the CYP2D6 gene may be the basis for designing a diagnostic method to determine those individuals who will or will not respond to the treatment, or alternatively, will respond at a lower level and thus may require more treatment, i.e., a greater dose of a drug. The diagnostic method may take one of several forms: for example, a direct DNA test (i.e., genotyping or haplotyping one or more of the polymorphic sites in the CYP2D6 gene), a serological test, or a physical exam measurement. The only requirement is that there be a good correlation between the diagnostic test results and the underlying CYP2D6 genotype or haplotype that is in turn correlated with the clinical response. In a preferred embodiment, this diagnostic method uses the predictive haplotyping method described above.

In another embodiment, the invention provides an isolated polynucleotide comprising a polymorphic variant of the CYP2D6 gene or a fragment of the gene which contains at least one of the novel polymorphic sites described herein. The nucleotide sequence of a variant CYP2D6 gene is identical to the reference genomic sequence for those portions of the gene examined, as described in the Examples below, except that it comprises a different nucleotide at one or more of the novel polymorphic sites PS1, PS2, PS3, PS4, PS6, PS10, PS12, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS22, PS23, PS24, PS26, PS27, PS28, PS29, PS32, PS34, PS36, PS37, PS38, PS39, PS40 and PS42, and may also comprise one or more additional polymorphisms selected from the group consisting of adenine at PS5, adenine at PS7, adenine at PS8, thymine at PS9, guanine at PS11, adenine at PS13, thymine at PS21, adenine at PS25, cytosine at PS30, guanine at PS31, adenine at PS33, cytosine at PS35 and cytosine at PS41. Similarly, the nucleotide sequence of a variant fragment

of the CYP2D6 gene is identical to the corresponding portion of the reference sequence except for having a different nucleotide at one or more of the novel polymorphic sites described herein. Thus, the invention specifically does not include polynucleotides comprising a nucleotide sequence identical to the reference sequence of the CYP2D6 gene (or other reported CYP2D6 sequences) or to portions of the reference sequence (or other reported CYP2D6 sequences), except for the haplotyping and genotyping oligonucleotides described above.

The location of a polymorphism in a variant CYP2D6 gene or fragment is preferably identified by aligning its sequence against SEQ ID NO:1. The polymorphism is selected from the group consisting of adenine at PS1, cytosine at PS2, cytosine at PS3, guanine at PS4, cytosine at PS6, cytosine at PS10, adenine at PS12, guanine at PS14, guanine at PS15, cytosine at PS16, thymine at PS17, thymine at PS18, guanine at PS19, cytosine at PS20, guanine at PS22, guanine at PS23, thymine at PS24, cytosine at PS26, adenine at PS27, cytosine at PS28, adenine at PS29, adenine at PS32, adenine at PS34, adenine at PS36, thymine at PS37, cytosine at PS38, cytosine at PS39, adenine at PS40 and thymine at PS42. In a preferred embodiment, the polymorphic variant comprises a naturally-occurring isogene of the CYP2D6 gene which is defined by any one of haplotypes 1-34 shown in Table 4 below.

Polymorphic variants of the invention may be prepared by isolating a clone containing the CYP2D6 gene from a human genomic library. The clone may be sequenced to determine the identity of the nucleotides at the novel polymorphic sites described herein. Any particular variant or fragment thereof, that is claimed herein could be prepared from this clone by performing *in vitro* mutagenesis using procedures well-known in the art. Any particular CYP2D6 variant or fragment thereof may also be prepared using synthetic or semi-synthetic methods known in the art.

CYP2D6 isogenes, or fragments thereof, may be isolated using any method that allows separation of the two "copies" of the CYP2D6 gene present in an individual, which, as readily understood by the skilled artisan, may be the same allele or different alleles. Separation methods include targeted *in vivo* cloning (TIVC) in yeast as described in WO 98/01573, U.S. Patent No. 5,866,404, and U.S. Patent No. 5,972,614. Another method, which is described in U.S. Patent No. 5,972,614, uses an allele specific oligonucleotide in combination with primer extension and exonuclease degradation to generate hemizygous DNA targets. Yet other methods are single molecule dilution (SMD) as described in Ruaño et al., *Proc. Natl. Acad. Sci.* 87:6296-6300, 1990; and allele specific PCR (Ruaño et al., 1989, *supra*; Ruaño et al., 1991, *supra*; Michalatos-Beloin et al., *supra*).

The invention also provides CYP2D6 genome anthologies, which are collections of at least two CYP2D6 isogenes found in a given population.. The population may be any group of at least two individuals, including but not limited to a reference population, a population group, a family population, a clinical population, and a same gender population. A CYP2D6 genome anthology may comprise individual CYP2D6 isogenes stored in separate containers such as microtest tubes, separate wells of a microtitre plate and the like. Alternatively, two or more groups of the CYP2D6 isogenes in

the anthology may be stored in separate containers. Individual isogenes or groups of such isogenes in a genome anthology may be stored in any convenient and stable form, including but not limited to in buffered solutions, as DNA precipitates, freeze-dried preparations and the like. A preferred CYP2D6 genome anthology of the invention comprises a set of isogenes defined by the haplotypes shown in Table 4 below.

An isolated polynucleotide containing a polymorphic variant nucleotide sequence of the invention may be operably linked to one or more expression regulatory elements in a recombinant expression vector capable of being propagated and expressing the encoded CYP2D6 protein in a prokaryotic or a eukaryotic host cell. Examples of expression regulatory elements which may be used include, but are not limited to, the lac system, operator and promoter regions of phage lambda, yeast promoters, and promoters derived from vaccinia virus, adenovirus, retroviruses, or SV40. Other regulatory elements include, but are not limited to, appropriate leader sequences, termination codons, polyadenylation signals, and other sequences required for the appropriate transcription and subsequent translation of the nucleic acid sequence in a given host cell. Of course, the correct combinations of expression regulatory elements will depend on the host system used. In addition, it is understood that the expression vector contains any additional elements necessary for its transfer to and subsequent replication in the host cell. Examples of such elements include, but are not limited to, origins of replication and selectable markers. Such expression vectors are commercially available or are readily constructed using methods known to those in the art (e.g., F. Ausubel et al., 1987, in "Current Protocols in Molecular Biology", John Wiley and Sons, New York, New York). Host cells which may be used to express the variant CYP2D6 sequences of the invention include, but are not limited to, eukaryotic and mammalian cells, such as animal, plant, insect and yeast cells, and prokaryotic cells, such as E. coli, or algal cells as known in the art. The recombinant expression vector may be introduced into the host cell using any method known to those in the art including, but not limited to, microinjection, electroporation, particle bombardment, transduction, and transfection using DEAE-dextran, lipofection, or calcium phosphate (see e.g., Sambrook et al. (1989) in "Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Press, Plainview, New York). In a preferred aspect, eukaryotic expression vectors that function in eukaryotic cells, and preferably mammalian cells, are used. Non-limiting examples of such vectors include vaccinia virus vectors, adenovirus vectors, herpes virus vectors, and baculovirus transfer vectors. Preferred eukaryotic cell lines include COS cells, CHO cells, HeLa cells, NIH/3T3 cells, and embryonic stem cells (Thomson, J. A. et al., 1998 *Science* 282:1145-1147). Particularly preferred host cells are mammalian cells.

As will be readily recognized by the skilled artisan, expression of polymorphic variants of the CYP2D6 gene will produce CYP2D6 mRNAs varying from each other at any polymorphic site retained in the spliced and processed mRNA molecules. These mRNAs can be used for the preparation of a CYP2D6 cDNA comprising a nucleotide sequence which is a polymorphic variant of the CYP2D6 reference coding sequence shown in Figure 2. Thus, the invention also provides

CYP2D6 mRNAs and corresponding cDNAs which comprise a nucleotide sequence that is identical to SEQ ID NO:2 (Fig. 2) (or its corresponding RNA sequence) for those regions of SEQ ID NO:2 that correspond to the examined portions of the CYP2D6 gene (as described in the Examples below), except for having one or more polymorphisms selected from the group consisting of adenine at a position corresponding to nucleotide 263, guanine at a position corresponding to nucleotide 281, guanine at a position corresponding to nucleotide 294, cytosine at a position corresponding to nucleotide 311, thymine at a position corresponding to nucleotide 319, thymine at a position corresponding to nucleotide 320, guanine at a position corresponding to nucleotide 325, cytosine at a position corresponding to nucleotide 333, adenine at a position corresponding to nucleotide 358, cytosine at a position corresponding to nucleotide 382, adenine at a position corresponding to nucleotide 406, adenine at a position corresponding to nucleotide 463, adenine at a position corresponding to nucleotide 1012, thymine at a position corresponding to nucleotide 1030, cytosine at a position corresponding to nucleotide 1083 and thymine at a position corresponding to nucleotide 1489, and may also comprise one or more additional polymorphisms selected from the group consisting of adenine at a position corresponding to nucleotide 19, adenine at a position corresponding to nucleotide 31, thymine at a position corresponding to nucleotide 100, adenine at a position corresponding to nucleotide 271, thymine at a position corresponding to nucleotide 336, cytosine at a position corresponding to nucleotide 408, guanine at a position corresponding to nucleotide 451, cytosine at a position corresponding to nucleotide 696 and cytosine at a position corresponding to nucleotide 1457. A particularly preferred polymorphic cDNA variant comprises the coding sequence of a CYP2D6 isogene defined by any one of haplotypes 1-10, 13-16, 18, 19, 21, 22, and 24-34. Fragments of these variant mRNAs and cDNAs are included in the scope of the invention, provided they contain one or more of the novel polymorphisms described herein. The invention specifically excludes polynucleotides identical to previously identified CYP2D6 mRNAs, cDNAs, or previously described fragments thereof. Polynucleotides comprising a variant CYP2D6 RNA or DNA sequence may be isolated from a biological sample using well-known molecular biological procedures or may be chemically synthesized.

As used herein, a polymorphic variant of a CYP2D6 gene, mRNA or cDNA fragment comprises at least one novel polymorphism identified herein and has a length of at least 10 nucleotides and may range up to the full length of the gene. Preferably, such fragments are between 100 and 3000 nucleotides in length, and more preferably between 200 and 2000 nucleotides in length, and most preferably between 500 and 1000 nucleotides in length.

In describing the CYP2D6 polymorphic sites identified herein, reference is made to the sense strand of the gene for convenience. However, as recognized by the skilled artisan, nucleic acid molecules containing the CYP2D6 gene or cDNA may be complementary double stranded molecules and thus reference to a particular site on the sense strand refers as well to the corresponding site on the complementary antisense strand. Thus, reference may be made to the same polymorphic site on either

strand and an oligonucleotide may be designed to hybridize specifically to either strand at a target region containing the polymorphic site. Thus, the invention also includes single-stranded polynucleotides which are complementary to the sense strand of the CYP2D6 genomic, mRNA and cDNA variants described herein.

Polynucleotides comprising a polymorphic gene variant or fragment of the invention may be useful for therapeutic purposes. For example, where a patient could benefit from expression, or increased expression, of a particular CYP2D6 protein isoform, an expression vector encoding the isoform may be administered to the patient. The patient may be one who lacks the CYP2D6 isogene encoding that isoform or may already have at least one copy of that isogene.

In other situations, it may be desirable to decrease or block expression of a particular CYP2D6 isogene. Expression of a CYP2D6 isogene may be turned off by transforming a targeted organ, tissue or cell population with an expression vector that expresses high levels of untranslatable mRNA or antisense RNA for the isogene or fragment thereof. Alternatively, oligonucleotides directed against the regulatory regions (e.g., promoter, introns, enhancers, 3′ untranslated region) of the isogene may block transcription. Oligonucleotides targeting the transcription initiation site, e.g., between positions −10 and +10 from the start site are preferred. Similarly, inhibition of transcription can be achieved using oligonucleotides that base-pair with region(s) of the isogene DNA to form triplex DNA (see e.g., Gee et al. in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y., 1994). Antisense oligonucleotides may also be designed to block translation of CYP2D6 mRNA transcribed from a particular isogene. It is also contemplated that ribozymes may be designed that can catalyze the specific cleavage of CYP2D6 mRNA transcribed from a particular isogene.

The untranslated mRNA, antisense RNA or antisense oligonucleotides may be delivered to a target cell or tissue by expression from a vector introduced into the cell or tissue *in vivo* or *ex vivo*. Alternatively, such molecules may be formulated as a pharmaceutical composition for administration to the patient. Oligoribonucleotides and/or oligodeoxynucleotides intended for use as antisense oligonucleotides may be modified to increase stability and half-life. Possible modifications include, but are not limited to phosphorothioate or 2′ O-methyl linkages, and the inclusion of nontraditional bases such as inosine and queosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytosine, guanine, thymine, and uracil which are not as easily recognized by endogenous nucleases.

The invention also provides an isolated polypeptide comprising a polymorphic variant of (a) the reference CYP2D6 amino acid sequence shown in Figure 3 or (b) a fragment of this reference sequence. The location of a variant amino acid in a CYP2D6 polypeptide or fragment of the invention is identified by aligning its sequence against SEQ ID NO:3 (Fig. 3). A CYP2D6 protein variant of the invention comprises an amino acid sequence identical to SEQ ID NO:3 for those regions of SEQ ID NO:3 that are encoded by examined portions of the CYP2D6 gene (as described in the Examples

below), except for having one or more variant amino acids selected from the group consisting ofhistidine at a position corresponding to amino acid position 88, arginine at a position corresponding to amino acid position 94, alanine at a position corresponding to amino acid position 104, phenylalanine at a position corresponding to amino acid position 107, phenylalanine at a position corresponding to amino acid position 107, valine at a position corresponding to amino acid position 109, isoleucine at a position corresponding to amino acid position 120, arginine at a position corresponding to amino acid position 128, isoleucine at a position corresponding to amino acid position 136, lysine at a position corresponding to amino acid position 155, methionine at a position corresponding to amino acid position 338, termination codon at a position corresponding to amino acid position 344 and cysteine at a position corresponding to amino acid position 497, and may also comprise one or more additional variant amino acids selected from the group consisting of methionine at a position corresponding to amino acid position 7, methionine at a position corresponding to amino acid position 11, serine at a position corresponding to amino acid position 34, methionine at a position corresponding to amino acid position 91, isoleucine at a position corresponding to amino acid position 136, glutamic acid at a position corresponding to amino acid position 151 and threonine at a position corresponding to amino acid position 486. Thus, a CYP2D6 fragment of the invention, also referred to herein as a CYP2D6 peptide variant, is any fragment of a CYP2D6 protein variant that contains one or more of the amino acid variations described herein. The invention specifically excludes amino acid sequences identical to those previously identified for CYP2D6, including SEQ ID NO:3, and previously described fragments thereof. CYP2D6 protein variants included within the invention comprise all amino acid sequences based on SEQ ID NO:3 and having any combination of amino acid variations described herein. In preferred embodiments, a CYP2D6 protein variant of the invention is encoded by an isogene defined by one of the observed haplotypes, 1-10, 13-16, 18, 19, 21, 22, and 24-34, shown in Table 4.

A CYP2D6 peptide variant of the invention is at least 6 amino acids in length and is preferably any number between 6 and 30 amino acids long, more preferably between 10 and 25, and most preferably between 15 and 20 amino acids long. Such CYP2D6 peptide variants may be useful as antigens to generate antibodies specific for one of the above CYP2D6 isoforms. In addition, the CYP2D6 peptide variants may be useful in drug screening assays.

A CYP2D6 variant protein or peptide of the invention may be prepared by chemical synthesis or by expressing an appropriate variant CYP2D6 genomic or cDNA sequence described above. Alternatively, the CYP2D6 protein variant may be isolated from a biological sample of an individual having a CYP2D6 isogene which encodes the variant protein. Where the sample contains two different CYP2D6 isoforms (i.e., the individual has different CYP2D6 isogenes), a particular CYP2D6 isoform of the invention can be isolated by immunoaffinity chromatography using an antibody which specifically binds to that particular CYP2D6 isoform but does not bind to the other CYP2D6 isoform.

The expressed or isolated CYP2D6 protein or peptide may be detected by methods known in

the art, including Coomassie blue staining, silver staining, and Western blot analysis using antibodies specific for the isoform of the CYP2D6 protein or peptide as discussed further below. CYP2D6 variant proteins and peptides can be purified by standard protein purification procedures known in the art, including differential precipitation, molecular sieve chromatography, ion-exchange chromatography, isoelectric focusing, gel electrophoresis, affinity and immunoaffinity chromatography and the like. (Ausubel et. al., 1987, In Current Protocols in Molecular Biology John Wiley and Sons, New York, New York). In the case of immunoaffinity chromatography, antibodies specific for a particular polymorphic variant may be used.

A polymorphic variant CYP2D6 gene of the invention may also be fused in frame with a heterologous sequence to encode a chimeric CYP2D6 protein. The non-CYP2D6 portion of the chimeric protein may be recognized by a commercially available antibody. In addition, the chimeric protein may also be engineered to contain a cleavage site located between the CYP2D6 and non-CYP2D6 portions so that the CYP2D6 protein may be cleaved and purified away from the non-CYP2D6 portion.

An additional embodiment of the invention relates to using a novel CYP2D6 protein isoform, or a fragment thereof, in any of a variety of drug screening assays. Such screening assays may be performed to identify agents that bind specifically to all known CYP2D6 protein isoforms or to only a subset of one or more of these isoforms. The agents may be from chemical compound libraries, peptide libraries and the like. The CYP2D6 protein or peptide variant may be free in solution or affixed to a solid support. In one embodiment, high throughput screening of compounds for binding to a CYP2D6 variant may be accomplished using the method described in PCT application WO84/03565, in which large numbers of test compounds are synthesized on a solid substrate, such as plastic pins or some other surface, contacted with the CYP2D6 protein(s) of interest and then washed. Bound CYP2D6 protein(s) are then detected using methods well-known in the art.

In another embodiment, a novel CYP2D6 protein isoform may be used in assays to measure the binding affinities of one or more candidate drugs targeting the CYP2D6 protein or to measure the enzymatic activity of CYP2D6 when using one or more candidate drugs as substrates.

In yet another embodiment, when a particular CYP2D6 haplotype or group of CYP2D6 haplotypes encodes a CYP2D6 protein variant with an amino acid sequence distinct from that of CYP2D6 protein isoforms encoded by other CYP2D6 haplotypes, then detection of that particular CYP2D6 haplotype or group of CYP2D6 haplotypes may be accomplished by detecting expression of the encoded CYP2D6 protein variant using any of the methods described herein or otherwise commonly known to the skilled artisan.

In another embodiment, the invention provides antibodies specific for and immunoreactive with one or more of the novel CYP2D6 variant proteins described herein. The antibodies may be either monoclonal or polyclonal in origin. The CYP2D6 protein or peptide variant used to generate the antibodies may be from natural or recombinant sources or produced by chemical synthesis using

synthesis techniques known in the art. If the CYP2D6 protein variant is of insufficient size to be antigenic, it may be conjugated, complexed, or otherwise covalently linked to a carrier molecule to enhance the antigenicity of the peptide. Examples of carrier molecules, include, but are not limited to, albumins (e.g., human, bovine, fish, ovine), and keyhole limpet hemocyanin (Basic and Clinical Immunology, 1991, Eds. D.P. Stites, and A.I. Terr, Appleton and Lange, Norwalk Connecticut, San Mateo, California).

In one embodiment, an antibody specifically immunoreactive with one of the novel protein isoforms described herein is administered to an individual to neutralize activity of the CYP2D6 isoform expressed by that individual. The antibody may be formulated as a pharmaceutical composition which includes a pharmaceutically acceptable carrier.

Antibodies specific for and immunoreactive with one of the novel protein isoforms described herein may be used to immunoprecipitate the CYP2D6 protein variant from solution as well as react with CYP2D6 protein isoforms on Western or immunoblots of polyacrylamide gels on membrane supports or substrates. In another preferred embodiment, the antibodies will detect CYP2D6 protein isoforms in paraffin or frozen tissue sections, or in cells which have been fixed or unfixed and prepared on slides, coverslips, or the like, for use in immunocytochemical, immunohistochemical, and immunofluorescence techniques.

In another embodiment, an antibody specifically immunoreactive with one of the novel CYP2D6 protein variants described herein is used in immunoassays to detect this variant in biological samples. In this method, an antibody of the present invention is contacted with a biological sample and the formation of a complex between the CYP2D6 protein variant and the antibody is detected. As described, suitable immunoassays include radioimmunoassay, Western blot assay, immunofluorescent assay, enzyme linked immunoassay (ELISA), chemiluminescent assay, immunohistochemical assay, immunocytochemical assay, and the like (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Press, New York, New York; Current Protocols in Molecular Biology, 1987, Eds. Ausubel et al., John Wiley and Sons, New York, New York). Standard techniques known in the art for ELISA are described in Methods in Immunodiagnosis, 2nd Ed., Eds. Rose and Bigazzi, John Wiley and Sons, New York 1980; and Campbell et al., 1984, Methods in Immunology, W.A. Benjamin, Inc.). Such assays may be direct, indirect, competitive, or noncompetitive as described in the art (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Pres, NY, NY; and Oellirich, M., 1984, J. Clin. Chem. Clin. Biochem., 22:895-904). Proteins may be isolated from test specimens and biological samples by conventional methods, as described in Current Protocols in Molecular Biology, supra.

Exemplary antibody molecules for use in the detection and therapy methods of the present invention are intact immunoglobulin molecules, substantially intact immunoglobulin molecules, or those portions of immunoglobulin molecules that contain the antigen binding site. Polyclonal or

monoclonal antibodies may be produced by methods conventionally known in the art (e.g., Kohler and Milstein, 1975, Nature, 256:495-497; Campbell Monoclonal Antibody Technology, the Production and Characterization of Rodent and Human Hybridomas, 1985, In: Laboratory Techniques in Biochemistry and Molecular Biology, Eds. Burdon et al., Volume 13, Elsevier Science Publishers, Amsterdam). The antibodies or antigen binding fragments thereof may also be produced by genetic engineering. The technology for expression of both heavy and light chain genes in E. coli is the subject of PCT patent applications, publication number WO 901443, WO 901443 and WO 9014424 and in Huse et al., 1989, Science, 246:1275-1281. The antibodies may also be humanized (e.g., Queen, C. et al. 1989 Proc. Natl. Acad. Sci.USA 86;10029).

Effect(s) of the polymorphisms identified herein on expression of CYP2D6 may be investigated by various means known in the art, such as by *in vitro* translation of mRNA transcripts of the CYP2D6 gene, cDNA orfragment thereof, or by preparing recombinant cells and/or nonhuman recombinant organisms, preferably recombinant animals, containing a polymorphic variant of the CYP2D6 gene. As used herein, "expression" includes but is not limited to one or more of the following: transcription of the gene into precursor mRNA; splicing and other processing of the precursor mRNA to produce mature mRNA; mRNA stability; translation of the mature mRNA(s) into CYP2D6 protein(s) (including effects of polymorphihsms on codon usage and tRNA availability); and glycosylation and/or other modifications of the translation product, if required for proper expression and function.

To prepare a recombinant cell of the invention, the desired CYP2D6 isogene, cDNA or coding sequence may be introduced into the cell in a vector such that the isogene, cDNA or coding sequence remains extrachromosomal. In such a situation, the gene will be expressed by the cell from the extrachromosomal location. In a preferred embodiment, the CYP2D6 isogene, cDNA or coding sequence is introduced into a cell in such a way that it recombines with the endogenous CYP2D6 gene present in the cell. Such recombination requires the occurrence of a double recombination event, thereby resulting in the desired CYP2D6 gene polymorphism. Vectors for the introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector or vector construct may be used in the invention. Methods such as electroporation, particle bombardment, calcium phosphate co-precipitation and viral transduction for introducing DNA into cells are known in the art; therefore, the choice of method may lie with the competence and preference of the skilled practitioner. Examples of cells into which the CYP2D6 isogene, cDNA or coding sequence may be introduced include, but are not limited to, continuous culture cells, such as COS, CHO, NIH/3T3, and primary or culture cells of the relevant tissue type, i.e., they express the CYP2D6 isogene, cDNA or coding sequence. Such recombinant cells can be used to compare the biological activities of the different protein variants.

Recombinant nonhuman organisms, i.e., transgenic animals, expressing a variant CYP2D6 gene, cDNA or coding sequence are prepared using standard procedures known in the art. Preferably,

a construct comprising the variant gene, cDNA or coding sequence is introduced into a nonhuman animal or an ancestor of the animal at an embryonic stage, i.e., the one-cell stage, or generally not later than about the eight-cell stage. Transgenic animals carrying the constructs of the invention can be made by several methods known to those having skill in the art. One method involves transfecting into the embryo a retrovirus constructed to contain one or more insulator elements, a gene or genes (or cDNA or coding sequence) of interest, and other components known to those skilled in the art to provide a complete shuttle vector harboring the insulated gene(s) as a transgene, see e.g., U.S. Patent No. 5,610,053. Another method involves directly injecting a transgene into the embryo. A third method involves the use of embryonic stem cells. Examples of animals into which the CYP2D6 isogene, cDNA or coding sequences may be introduced include, but are not limited to, mice, rats, other rodents, and nonhuman primates (see "The Introduction of Foreign Genes into Mice" and the cited references therein, In: Recombinant DNA, Eds. J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller; W.H. Freeman and Company, New York, pages 254-272). Transgenic animals stably expressing a human CYP2D6 isogene, cDNA or coding sequence and producing the encoded human CYP2D6 protein can be used as biological models for studying diseases related to abnormal CYP2D6 expression and/or activity, and for screening and assaying various candidate drugs, compounds, and treatment regimens to reduce the symptoms or effects of these diseases.

An additional embodiment of the invention relates to pharmaceutical compositions for treating disorders affected by expression or function of a novel CYP2D6 isogene described herein. The pharmaceutical composition may comprise any of the following active ingredients: a polynucleotide comprising one of these novel CYP2D6 isogenes (or cDNAs or coding sequences); an antisense oligonucleotide directed against one of the novel CYP2D6 isogenes, a polynucleotide encoding such an antisense oligonucleotide, or another compound which inhibits expression of a novel CYP2D6 isogene described herein. Preferably, the composition contains the active ingredient in a therapeutically effective amount. By therapeutically effective amount is meant that one or more of the symptoms relating to disorders affected by expression or function of a novel CYP2D6 isogene is reduced and/or eliminated. The composition also comprises a pharmaceutically acceptable carrier, examples of which include, but are not limited to, saline, buffered saline, dextrose, and water. Those skilled in the art may employ a formulation most suitable for the active ingredient, whether it is a polynucleotide, oligonucleotide, protein, peptide or small molecule antagonist. The pharmaceutical composition may be administered alone or in combination with at least one other agent, such as a stabilizing compound. Administration of the pharmaceutical composition may be by any number of routes including, but not limited to oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, intradermal, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

For any composition, determination of the therapeutically effective dose of active ingredient and/or the appropriate route of administration is well within the capability of those skilled in the art. For example, the dose can be estimated initially either in cell culture assays or in animal models. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. The exact dosage will be determined by the practitioner, in light of factors relating to the patient requiring treatment, including but not limited to severity of the disease state, general health, age, weight and gender of the patient, diet, time and frequency of administration, other drugs being taken by the patient, and tolerance/response to the treatment.

Any or all analytical and mathematical operations involved in practicing the methods of the present invention may be implemented by a computer. In addition, the computer may execute a program that generates views (or screens) displayed on a display device and with which the user can interact to view and analyze large amounts of information relating to the CYP2D6 gene and its genomic variation, including chromosome location, gene structure, and gene family, gene expression data, polymorphism data, genetic sequence data, and clinical data population data (e.g., data on ethnogeographic origin, clinical responses, genotypes, and haplotypes for one or more populations). The CYP2D6 polymorphism data described herein may be stored as part of a relational database (e.g., an instance of an Oracle database or a set of ASCII flat files). These polymorphism data may be stored on the computer's hard drive or may, for example, be stored on a CD-ROM or on one or more other storage devices accessible by the computer. For example, the data may be stored on one or more databases in communication with the computer via a network.

Preferred embodiments of the invention are described in the following examples. Other embodiments within the scope of the claims herein will be apparent to one skilled in the art from consideration of the specification or practice of the invention as disclosed herein. It is intended that the specification, together with the examples, be considered exemplary only, with the scope and spirit of the invention being indicated by the claims which follow the examples.

EXAMPLES

The Examples herein are meant to exemplify the various aspects of carrying out the invention and are not intended to limit the scope of the invention in any way. The Examples do not include detailed descriptions for conventional methods employed, such as in the performance of genomic DNA isolation, PCR and sequencing procedures. Such methods are well-known to those skilled in the art and are described in numerous publications, for example, Sambrook, Fritsch, and Maniatis, "Molecular Cloning: A Laboratory Manual", 2nd Edition, Cold Spring Harbor Laboratory Press, USA, (1989).

EXAMPLE 1

This example illustrates examination of various regions of the CYP2D6 gene for polymorphic sites.

Amplification of Target Regions

The following target regions were amplified using either the PCR primers represented below or 'tailed' PCR primers, each of which includes a universal sequence forming a noncomplementary 'tail' attached to the 5′ end of each unique sequence in the PCR primer pairs. The universal 'tail' sequence for the forward PCR primers comprises the sequence 5′-TGTAAAACGACGGCCAGT-3′ (SEQ ID NO:149) and the universal 'tail' sequence for the reverse PCR primers comprises the sequence 5′-AGGAAACAGCTATGACCAT-3′ (SEQ ID NO:150). The nucleotide positions of the first and last nucleotide of the forward and reverse primers for each region amplified are presented below and correspond to positions in SEQ ID NO:1 (Figure 1).

PCR Primer Pairs

Fragment No.	Forward Primer	Reverse Primer	PCR Product
Fragment 1	378-400	complement of 934-913	557 nt
Fragment 2	645-667	complement of 1184-1163	540 nt
Fragment 3	833-855	complement of 1363-1343	531 nt
Fragment 4	1701-1723	complement of 2203-2182	503 nt
Fragment 5	2342-2364	complement of 2942-2922	601 nt
Fragment 6	2652-2671	complement of 3244-3220	593 nt
Fragment 7	3213-3235	complement of 3794-3773	582 nt
Fragment 8	3604-3627	complement of 4184-4164	581 nt
Fragment 9	4008-4028	complement of 4555-4533	548 nt
Fragment 10	4651-4673	complement of 5094-5072	444 nt
Fragment 11	4867-4886	complement of 5297-5275	431 nt
Fragment 12	4968-4987	complement of 5435-5413	468 nt

These primer pairs were used in PCR reactions containing genomic DNA isolated from immortalized cell lines for each member of the Index Repository. The PCR reactions were carried out under the following conditions:

Reaction volume	= 10 µl
10 x Advantage 2 Polymerase reaction buffer (Clontech)	= 1 µl
100 ng of human genomic DNA	= 1 µl
10 mM dNTP	= 0.4 µl
Advantage 2 Polymerase enzyme mix (Clontech)	= 0.2 µl
Forward Primer (10 µM)	= 0.4 µl
Reverse Primer (10 µM)	= 0.4 µl
Water	= 6.6µl

Amplification profile:
97°C - 2 min. 1 cycle

97°C - 15 sec.
70°C - 45 sec.
72°C - 45 sec. } 10 cycles

97°C - 15 sec.
64°C - 45 sec.
72°C - 45 sec. } 35 cycles

Sequencing of PCR Products

The PCR products were purified using a Whatman/Polyfiltronics 100 µl 384 well unifilter plate essentially according to the manufacturers protocol. The purified DNA was eluted in 50 µl of distilled water. Sequencing reactions were set up using Applied Biosystems Big Dye Terminator chemistry essentially according to the manufacturers protocol. The purified PCR products were sequenced in both directions using either the primer sets represented below with the positions of their first and last nucleotide corresponding to positions in Figure 1, or the appropriate universal 'tail' sequence as a primer. Reaction products were purified by isopropanol precipitation, and run on an Applied Biosystems 3700 DNA Analyzer.

Sequencing Primer Pairs

Fragment No.	Forward Primer	Reverse Primer
Fragment 1	Tailed sequence	
Fragment 2	778-797	complement of 1155-1136
Fragment 3	Tailed sequence	
Fragment 4	Tailed sequence	
Fragment 5	2471-2490	complement of 2890-2871
Fragment 6	2687-2706	complement of 3133-3113
Fragment 7	Tailed sequence	
Fragment 8	3688-3707	complement of 4079-4058
Fragment 9	4036-4055	complement of 4469-4450
Fragment 10	Tailed sequence	
Fragment 11	4915-4934	complement of 5260-5241
Fragment 12	5070-5089	complement of 5412-5392

Analysis of Sequences for Polymorphic Sites

Sequence information for a minimum of 80 humans was analyzed for the presence of polymorphisms using the Polyphred program (Nickerson et al., *Nucleic Acids Res.* 14:2745-2751, 1997). The presence of a polymorphism was confirmed on both strands. The polymorphisms and their locations in the CYP2D6 reference genomic sequence (SEQ ID NO:1) are listed in Table 2 below.

Table 2. Polymorphic Sites Identified in the CYP2D6 Gene

Polymorphic Site Number	PolyId(a)	Nucleotide Position	Reference Allele	Variant Allele	CDS Variant Position	AA Variant
PS1	3578169	636	G	A		
PS2	3578173	678	T	C		
PS3	3578177	769	G	C		
PS4	3578179	776	A	G		
PS5(R)	3578181	825	G	A		
PS6	3578183	915	T	C		
PS7(R)	3578185	1019	G	A	19	V7M
PS8(R)	3578187	1031	G	A	31	V11M
PS9(R)	3578189	1100	C	T	100	P34S
PS10	8552536	1827	G	C		
PS11(R)	6823986	1843	T	G		
PS12	8552444	1966	G	A	263	R88H
PS13(R)	8060758	1974	C	A	271	L91M
PS14	8060761	1984	A	G	281	H94R
PS15	8060764	1997	C	G	294	T98T
PS16	8552352	2014	T	C	311	V104A
PS17	8552259	2022	A	T	319	T107F
PS18	8060767	2023	C	T	320	T107F
PS19	8552166	2028	A	G	325	I109V
PS20	8552073	2036	T	C	333	G111G
PS21(R)	8060770	2039	C	T	336	F112F
PS22	8551980	2062	A	G		
PS23	8551887	2067	T	G		
PS24	8551795	2118	C	T		
PS25(R)	8551703	2170	G	A		
PS26	8551610	2179	G	C		
PS27	3578191	2611	T	A	358	F120I
PS28	3578193	2635	T	C	382	W128R
PS29	3578195	2659	G	A	406	V136I
PS30(R)	3578199	2661	G	C	408	V136I
PS31(R)	3578201	2704	C	G	451	Q151E
PS32	3578203	2716	G	A	463	E155K
PS33(R)	3578210	2846	G	A		
PS34	3578212	3292	G	A		
PS35(R)	3578214	3470	T	C	696	H232H
PS36	3578221	4183	G	A	1012	V338M
PS37	3578223	4201	C	T	1030	R344!
PS38	3578229	4254	T	C	1083	H361H
PS39	3578235	4384	A	C		
PS40	3578241	4435	C	A		
PS41(R)	3578249	5180	G	C	1457	S486T
PS42	3578251	5212	C	T	1489	R497C

(a)PolyId is a unique identifier assigned to each PS by Genaissance Pharmaceuticals, Inc.

(R)Reported previously

EXAMPLE 2

This example illustrates analysis of the CYP2D6 polymorphisms identified in the Index Repository for human genotypes and haplotypes.

The different genotypes containing these polymorphisms that were observed in unrelated members of the reference population are shown in Table 3 below, with the haplotype pair indicating the combination of haplotypes determined for the individual using the haplotype derivation protocol described below. In Table 3, homozygous positions are indicated by one nucleotide and heterozygous positions are indicated by two nucleotides. Missing nucleotides in any given genotype in Table 3 were inferred based on linkage disequilibrium and/or Mendelian inheritance.

Table 3 (Part 1). Genotypes and Haplotype Pairs Observed for CYP2D6 Gene

Genotype Number	HAP Pair		Polymorphic Sites PS1	PS2	PS3	PS4	PS5	PS6	PS7	PS8	PS9	PS10
1	15	15	G	T	G	A	G	T	G	G	C	G
2	19	19	G	T	G	A	G	T	G	G	C	G
3	32	32	G	T	G	A	G	T	G	G	T	G
4	7	7	G	T	G	A	G	T	G	A	C	G
5	15	20	G	T	G	A	G	T	G	G	C	G
6	15	19	G	T	G	A	G	T	G	G	C	G
7	23	33	G	T	G	A	G	T	G	G	C/T	G
8	22	20	G	T	G	A	G	T	G	G	C	G
9	23	24	G	T	G	A	G	T	G	G	C	G
10	23	25	G	T	G	A	G	T	G	G	C/T	G
11	15	9	G	T	G	A	G	T	G	G	C	G/C
12	32	21	G	T	G	A	G	T	G	G	T/C	G
13	23	16	G	T	G	A	G	T	G	G	C	G
14	20	28	G	T	G	A	G	T	G	G	C/T	G
15	32	13	G	T	G	A	G	T	G	G	T/C	G
16	32	26	G	T	G	A	G	T	G	G	T	G
17	23	1	G/A	T	G	A	G	T	G	G	C	G
18	32	20	G	T	G	A	G	T	G	G	T/C	G
19	15	12	G	T	G	A	G	T	G	G	C	G
20	23	26	G	T	G	A	G	T	G	G	C/T	G
21	23	18	G	T	G	A	G	T	G	G	C	G
22	23	27	G	T	G	A	G	T	G	G	C/T	G
23	32	2	G	T/C	G	A	G	T	G	G	T/C	G
24	15	34	G	T	G	A	G	T	G	G	C/T	G
25	15	14	G	T	G	A	G	T	G	G	C	G
26	23	8	G	T	G	A	G	T	G	G/A	C	G
27	23	28	G	T	G	A	G	T	G	G	C/T	G
28	15	7	G	T	G	A	G	T	G	G/A	C	G
29	23	10	G	T	G	A	G	T	G	G	C	G
30	23	5	G	T	G	A	G	T/C	G	G	C	G
31	32	17	G	T	G	A	G	T	G	G	T/C	G
32	15	26	G	T	G	A	G	T	G	G	C/T	G
33	15	11	G	T	G	A	G	T	G	G	C	G
34	32	4	G	T	G	A	G/A	T	G	G	T/C	G
35	26	19	G	T	G	A	G	T	G	G	T/C	G
36	23	4	G	T	G	A	G/A	T	G	G	C	G
37	32	12	G	T	G	A	G	T	G	G	T/C	G
38	30	29	G	T	G	A	G	T	G	G	T	G
39	23	6	G	T	G	A	G	T	G/A	G	C	G
40	23	32	G	T	G	A	G	T	G	G	C/T	G
41	23	15	G	T	G	A	G	T	G	G	C	G
42	15	3	G	T	G/C	A/G	G	T	G	G	C/T	G
43	26	2	G	T/C	G	A	G	T	G	G	T/C	G
44	15	32	G	T	G	A	G	T	G	G	C/T	G
45	23	22	G	T	G	A	G	T	G	G	C	G
46	4	31	G	T	G	A	A/G	T	G	G	C/T	G
47	15	31	G	T	G	A	G	T	G	G	C/T	G
48	15	4	G	T	G	A	G/A	T	G	G	C	G

Table 3 (Part 2). Genotypes and Haplotype Pairs Observed for CYP2D6 Gene

Genotype Number	HAP Pair		Polymorphic Sites PS11	PS12	PS13	PS14	PS15	PS16	PS17	PS18	PS19	PS20
1	15	15	G	G	C	A	C	T	A	C	A	T
2	19	19	G	G	C	A	C	T	A	T	A	T
3	32	32	G	G	C	A	C	T	A	C	A	T
4	7	7	G	G	C	A	C	T	A	C	A	T
5	15	20	G/T	G	C	A	C	T	A	C	A	T
6	15	19	G	G	C	A	C	T	A	C/T	A	T
7	23	33	T/G	G	C	A	C	T	A	C	A	T
8	22	20	T	G	C	A	C	T	A	C	A	T
9	23	24	T	G	C	A	C	T	A	C	A	T
10	23	25	T/G	G	C/A	A	C	T	A	C	A	T
11	15	9	G	G	C	A	C	T	A	C/T	A	T
12	32	21	G/T	G	C	A	C	T	A	C	A	T
13	23	16	T/G	G	C	A	C	T	A	C	A	T
14	20	28	T/G	G	C/A	A	C/G	T	A	C	A	T
15	32	13	G	G	C	A	C	T	A	C	A	T
16	32	26	G	G	C/A	A	C/G	T	A	C	A	T
17	23	1	T	G	C	A	C	T	A	C	A	T
18	32	20	G/T	G	C	A	C	T	A	C	A	T
19	15	12	G	G	C	A	C	T	A	C	A	T
20	23	26	T/G	G	C/A	A	C/G	T	A	C	A	T
21	23	18	T/G	G	C	A	C	T	A	C/T	A	T
22	23	27	T/G	G	C/A	A	C/G	T	A	C	A	T
23	32	2	G	G	C	A	C	T	A	C	A	T
24	15	34	G/T	G	C/A	A	C/G	T	A	C	A	T
25	15	14	G	G	C	A	C	T	A	C	A	T
26	23	8	T/G	G	C	A	C	T	A	C	A	T
27	23	28	T/G	G	C/A	A	C/G	T	A	C	A	T
28	15	7	G	G	C	A	C	T	A	C	A	T
29	23	10	T/G	G/A	C	A	C	T	A	C	A	T
30	23	5	T	G	C	A	C	T	A	C	A	T
31	32	17	G	G	C	A	C	T	A	C	A	T
32	15	26	G	G	C/A	A	C/G	T	A	C	A	T
33	15	11	G	G	C	A	C	T	A	C	A	T
34	32	4	G	G	C	A	C	T	A	C	A	T
35	26	19	G	G	A/C	A	G/C	T	A	C/T	A	T
36	23	4	T/G	G	C	A	C	T	A	C	A	T
37	32	12	G	G	C	A	C	T	A	C	A	T
38	30	29	G	G	A	G	G	T	A	C	A	T
39	23	6	T/G	G	C	A	C	T	A	C	A	T
40	23	32	T/G	G	C	A	C	T	A	C	A	T
41	23	15	T/G	G	C	A	C	T	A	C	A	T
42	15	3	G	G	C	A	C	T	A	C	A	T
43	26	2	G	G	A/C	A	G/C	T	A	C	A	T
44	15	32	G	G	C	A	C	T	A	C	A	T
45	23	22	T	G	C	A	C	T	A	C	A	T
46	4	31	G	G	C	A	C	T	A	C	A	T
47	15	31	G	G	C	A	C	T	A	C	A	T
48	15	4	G	G	C	A	C	T	A	C	A	T

Table 3 (Part 3). Genotypes and Haplotype Pairs Observed for CYP2D6 Gene

Genotype			Polymorphic Sites									
Number	HAP Pair		PS21	PS22	PS23	PS24	PS25	PS26	PS27	PS28	PS29	PS30
1	15	15	C	A	T	C	G	G	T	T	G	C
2	19	19	C	A	T	C	G	G	T	T	G	C
3	32	32	T	A	T	C	G	G	T	T	G	C
4	7	7	C	A	T	C	G	G	T	T	G	C
5	15	20	C	A	T	C	G/A	G	T	T	G	C/G
6	15	19	C	A	T	C	G	G	T	T	G	C
7	23	33	C/T	A	T	C	G	G	T	T	G	G/C
8	22	20	C	A	T	C	G/A	G	T	T	G	G
9	23	24	C	A	T	C/T	G	G	T	T	G	G
10	23	25	C	A	T	C	G	G	T	T	G	G/C
11	15	9	C	A	T	C	G	G	T	T/C	G	C
12	32	21	T/C	A	T	C	G	G	T	T	G	C
13	23	16	C	A	T	C	G	G	T	T	G	G
14	20	28	C	A	T	C	A/G	G	T	T	G	G/C
15	32	13	T/C	A	T	C	G	G	T/A	T	G	C
16	32	26	T/C	A	T	C	G	G	T	T	G	C
17	23	1	C	A	T	C	G	G	T	T	G	G
18	32	20	T/C	A	T	C	G/A	G	T	T	G	C/G
19	15	12	C	A	T/G	C	G	G	T	T	G	C/G
20	23	26	C	A	T	C	G	G	T	T	G	G/C
21	23	18	C	A	T	C	G	G	T	T	G	G/C
22	23	27	C	A	T	C	G	G	T	T	G	G/C
23	32	2	T/C	A	T	C	G	G	T	T	G	C
24	15	34	C	A	T	C	G	G	T	T	G	C
25	15	14	C	A	T	C	G	G	T	T	G	C
26	23	8	C	A	T	C	G	G	T	T	G	G
27	23	28	C	A	T	C	G	G	T	T	G	G/C
28	15	7	C	A	T	C	G	G	T	T	G	C
29	23	10	C	A	T	C	G	G	T	T	G	G/C
30	23	5	C	A	T	C	G	G	T	T	G	G
31	32	17	T/C	A	T	C	G	G	T	T	G	C/G
32	15	26	C	A	T	C	G	G	T	T	G	C
33	15	11	C	A	T/G	C	G	G/C	T	T	G	C/G
34	32	4	T/C	A	T	C	G	G	T	T	G/A	C
35	26	19	C	A	T	C	G	G	T	T	G	C
36	23	4	C	A	T	C	G	G	T	T	G/A	G/C
37	32	12	T/C	A	T/G	C	G	G	T	T	G	C/G
38	30	29	C	A	T	C	G	G	T	T	G	C
39	23	6	C	A	T	C	G	G	T	T	G	G/C
40	23	32	C/T	A	T	C	G	G	T	T	G	G/C
41	23	15	C	A	T	C	G	G	T	T	G	G/C
42	15	3	C/T	A	T	C	G	G	T	T	G	C
43	26	2	C	A	T	C	G	G	T	T	G	C
44	15	32	C/T	A	T	C	G	G	T	T	G	C
45	23	22	C	A	T	C	G	G	T	T	G	G
46	4	31	C/T	A	T	C	G	G	T	T	A/G	C
47	15	31	C/T	A	T	C	G	G	T	T	G	C
48	15	4	C	A	T	C	G	G	T	T	G/A	C

Table 3 (Part 4). Genotypes and Haplotype Pairs Observed for CYP2D6 Gene

Genotype Number	HAP Pair		Polymorphic Sites PS31	PS32	PS33	PS34	PS35	PS36	PS37	PS38	PS39	PS40
1	15	15	C	G	G	G	T	G	C	T	C	C
2	19	19	C	G	G	G	T	G	C	T	C	C
3	32	32	C	G	G	G	T	G	C	T	C	C
4	7	7	C	G	G	G	T	G	C	T	C	C
5	15	20	C	G	G	G	T	G	C	T	C/A	C
6	15	19	C	G	G	G	T	G	C	T	C	C
7	23	33	C	G	G	G	T	G	C/T	T	A/C	C
8	22	20	C	G	G	G	T	G	C	T	A	C
9	23	24	C	G	G	G	T	G	C	T	A	C
10	23	25	C	G	G	G	T	G	C	T	A/C	C
11	15	9	C	G	G	G	T	G	C	T	C	C
12	32	21	C	G	G	G	T	G	C	T	C/A	C
13	23	16	C	G	G	G	T	G	C	T	A	C
14	20	28	C	G	G	G	T	G	C	T	A/C	C
15	32	13	C	G	G	G	T	G	C	T	C	C
16	32	26	C	G	G/A	G	T	G	C	T	C	C
17	23	1	C	G	G	G	T	G	C	T	A	C
18	32	20	C	G	G	G	T	G	C	T	C/A	C
19	15	12	C	G	G	G	T	G	C	T	C/A	C
20	23	26	C	G	G/A	G	T	G	C	T	A/C	C
21	23	18	C	G	G	G	T/C	G	C	T	A/C	C
22	23	27	C	G	G	G	T	G	C	T	A/C	C/A
23	32	2	C	G	G	G	T	G	C	T	C	C
24	15	34	C	G	G/A	G	T	G	C	T	C/A	C
25	15	14	C	G/A	G	G	T	G	C	T	C	C
26	23	8	C	G	G	G	T	G	C	T	A/C	C
27	23	28	C	G	G	G	T	G	C	T	A/C	C
28	15	7	C	G	G	G	T	G	C	T	C	C
29	23	10	C	G	G	G	T	G	C	T	A/C	C
30	23	5	C	G	G	G	T	G	C	T	A	C
31	32	17	C	G	G	G	T	G	C	T	C/A	C
32	15	26	C	G	G/A	G	T	G	C	T	C	C
33	15	11	C	G	G	G	T	G	C	T	C/A	C
34	32	4	C	G	G	G/A	T	G/A	C	T	C	C
35	26	19	C	G	A/G	G	T	G	C	T	C	C
36	23	4	C	G	G	G/A	T	G/A	C	T	A/C	C
37	32	12	C	G	G	G	T	G	C	T	C/A	C
38	30	29	C	G	G/A	G	T	G	C	T	C	A/C
39	23	6	C/G	G	G	G	T	G	C	T	A/C	C
40	23	32	C	G	G	G	T	G	C	T	A/C	C
41	23	15	C	G	G	G	T	G	C	T	A/C	C
42	15	3	C	G	G	G	T	G	C	T	C	C
43	26	2	C	G	A/G	G	T	G	C	T	C	C
44	15	32	C	G	G	G	T	G	C	T	C	C
45	23	22	C	G	G	G	T	G	C	T	A	C
46	4	31	C	G	G/A	A/G	T	A/G	C	T	C	C
47	15	31	C	G	G/A	G	T	G	C	T	C	C
48	15	4	C	G	G	G/A	T	G/A	C	T	C	C

Table 3 (Part 5). Genotypes and Haplotype Pairs Observed for CYP2D6 Gene

Genotype Number	HAP Pair		Polymorphic Sites PS41	PS42
1	15	15	C	C
2	19	19	C	C
3	32	32	C	C
4	7	7	C	C
5	15	20	C/G	C
6	15	19	C	C
7	23	33	G/C	C
8	22	20	C/G	C
9	23	24	G/C	C
10	23	25	G/C	C
11	15	9	C	C
12	32	21	C/G	C
13	23	16	G/C	C
14	20	28	G/C	C
15	32	13	C	C
16	32	26	C	C
17	23	1	G/C	C
18	32	20	C/G	C
19	15	12	C/G	C
20	23	26	G/C	C
21	23	18	G/C	C
22	23	27	G/C	C
23	32	2	C	C
24	15	34	C/G	C
25	15	14	C	C
26	23	8	G/C	C
27	23	28	G/C	C
28	15	7	C	C
29	23	10	G/C	C
30	23	5	G/C	C
31	32	17	C/G	C
32	15	26	C	C
33	15	11	C/G	C
34	32	4	C	C
35	26	19	C	C
36	23	4	G/C	C
37	32	12	C/G	C
38	30	29	C	C
39	23	6	G/C	C
40	23	32	G/C	C
41	23	15	G/C	C
42	15	3	C	C
43	26	2	C	C
44	15	32	C	C
45	23	22	G/C	C
46	4	31	C	C
47	15	31	C	C
48	15	4	C	C

The haplotype pairs shown in Table 3 were estimated from the unphased genotypes using a computer-implemented extension of Clark's algorithm (Clark, A.G. 1990 *Mol Bio Evol* 7, 111-122) for assigning haplotypes to unrelated individuals in a population sample, as described in PCT/US01/12831, filed April 18, 2001. In this method, haplotypes are assigned directly from individuals who are homozygous at all sites or heterozygous at no more than one of the variable sites. This list of haplotypes is then used to deconvolute the unphased genotypes in the remaining (multiply heterozygous) individuals. In the present analysis, the list of haplotypes was augmented with haplotypes obtained from two families (one three-generation Caucasian family and one two-generation African-American family).

By following this protocol, it was determined that the Index Repository examined herein and, by extension, the general population contains the 34 human CYP2D6 haplotypes shown in Table 4 below.

A CYP2D6 isogene defined by a full-haplotype shown in Table 4 below comprises the regions of the SEQ ID NOS indicated in Table 4, with their corresponding set of polymorphic locations and identities, which are also set forth in Table 4.

Table 4 (Part 1). Haplotypes of the CYP2D6 gene.

Regions Examined(a)	PS No.(b)	PS Position(c)	Haplotype Number(d)									
			1	2	3	4	5	6	7	8	9	10
378-1363	1	636/30	A	G	G	G	G	G	G	G	G	G
378-1363	2	678/150	T	C	T	T	T	T	T	T	T	T
378-1363	3	769/270	G	G	C	G	G	G	G	G	G	G
378-1363	4	776/390	A	A	G	A	A	A	A	A	A	A
378-1363	5	825/510	G	G	G	A	G	G	G	G	G	G
378-1363	6	915/630	T	T	T	T	C	T	T	T	T	T
378-1363	7	1019/750	G	G	G	G	G	A	G	G	G	G
378-1363	8	1031/870	G	G	G	G	G	G	A	A	G	G
378-1363	9	1100/990	C	C	T	C	C	C	C	C	C	C
1701-2203	10	1827/1110	G	G	G	G	G	G	G	G	C	G
1701-2203	11	1843/1230	T	G	G	G	T	G	G	G	G	G
1701-2203	12	1966/1350	G	G	G	G	G	G	G	G	G	A
1701-2203	13	1974/1470	C	C	C	C	C	C	C	C	C	C
1701-2203	14	1984/1590	A	A	A	A	A	A	A	A	A	A
1701-2203	15	1997/1710	C	C	C	C	C	C	C	C	C	C
1701-2203	16	2014/1830	T	T	T	T	T	T	T	T	T	T
1701-2203	17	2022/1950	A	A	A	A	A	A	A	A	A	A
1701-2203	18	2023/2070	C	C	C	C	C	C	C	C	T	C
1701-2203	19	2028/2190	A	A	A	A	A	A	A	A	A	A
1701-2203	20	2036/2310	T	T	T	T	T	T	T	T	T	T
1701-2203	21	2039/2430	C	C	T	C	C	C	C	C	C	C
1701-2203	22	2062/2550	A	A	A	A	A	A	A	A	A	A
1701-2203	23	2067/2670	T	T	T	T	T	T	T	T	T	T
1701-2203	24	2118/2790	C	C	C	C	C	C	C	C	C	C
1701-2203	25	2170/2910	G	G	G	G	G	G	G	G	G	G
1701-2203	26	2179/3030	G	G	G	G	G	G	G	G	G	G
2342-4555	27	2611/3150	T	T	T	T	T	T	T	T	T	T
2342-4555	28	2635/3270	T	T	T	T	T	T	T	T	C	T
2342-4555	29	2659/3390	G	G	G	A	G	G	G	G	G	G
2342-4555	30	2661/3510	G	C	C	C	G	C	C	G	C	C
2342-4555	31	2704/3630	C	C	C	C	C	G	C	C	C	C
2342-4555	32	2716/3750	G	G	G	G	G	G	G	G	G	G
2342-4555	33	2846/3870	G	G	G	G	G	G	G	G	G	G
2342-4555	34	3292/3990	G	G	G	A	G	G	G	G	G	G
2342-4555	35	3470/4110	T	T	T	T	T	T	T	T	T	T
2342-4555	36	4183/4230	G	G	G	A	G	G	G	G	G	G
2342-4555	37	4201/4350	C	C	C	C	C	C	C	C	C	C
2342-4555	38	4254/4470	T	T	T	T	T	T	T	T	T	T
2342-4555	39	4384/4590	A	C	C	C	A	C	C	C	C	C
2342-4555	40	4435/4710	C	C	C	C	C	C	C	C	C	C
4651-5435	41	5180/4830	C	C	C	C	C	C	C	C	C	C
4651-5435	42	5212/4950	C	C	C	C	C	C	C	C	C	C

Table 4 (Part 2). Haplotypes of the CYP2D6 gene.

Regions Examined(a)	PS No.(b)	PS Position(c)	Haplotype Number(d)									
			11	12	13	14	15	16	17	18	19	20
378-1363	1	636/30	G	G	G	G	G	G	G	G	G	G
378-1363	2	678/150	T	T	T	T	T	T	T	T	T	T
378-1363	3	769/270	G	G	G	G	G	G	G	G	G	G
378-1363	4	776/390	A	A	A	A	A	A	A	A	A	A
378-1363	5	825/510	G	G	G	G	G	G	G	G	G	G
378-1363	6	915/630	T	T	T	T	T	T	T	T	T	T
378-1363	7	1019/750	G	G	G	G	G	G	G	G	G	G
378-1363	8	1031/870	G	G	G	G	G	G	G	G	G	G
378-1363	9	1100/990	C	C	C	C	C	C	C	C	C	C
1701-2203	10	1827/1110	G	G	G	G	G	G	G	G	G	G
1701-2203	11	1843/1230	G	G	G	G	G	G	G	G	G	T
1701-2203	12	1966/1350	G	G	G	G	G	G	G	G	G	G
1701-2203	13	1974/1470	C	C	C	C	C	C	C	C	C	C
1701-2203	14	1984/1590	A	A	A	A	A	A	A	A	A	A
1701-2203	15	1997/1710	C	C	C	C	C	C	C	C	C	C
1701-2203	16	2014/1830	T	T	T	T	T	T	T	T	T	T
1701-2203	17	2022/1950	A	A	A	A	A	A	A	A	A	A
1701-2203	18	2023/2070	C	C	C	C	C	C	C	T	T	C
1701-2203	19	2028/2190	A	A	A	A	A	A	A	A	A	A
1701-2203	20	2036/2310	T	T	T	T	T	T	T	T	T	T
1701-2203	21	2039/2430	C	C	C	C	C	C	C	C	C	C
1701-2203	22	2062/2550	A	A	A	A	A	A	A	A	A	A
1701-2203	23	2067/2670	G	G	T	T	T	T	T	T	T	T
1701-2203	24	2118/2790	C	C	C	C	C	C	C	C	C	C
1701-2203	25	2170/2910	G	G	G	G	G	G	G	G	G	A
1701-2203	26	2179/3030	C	G	G	G	G	G	G	G	G	G
2342-4555	27	2611/3150	T	T	A	T	T	T	T	T	T	T
2342-4555	28	2635/3270	T	T	T	T	T	T	T	T	T	T
2342-4555	29	2659/3390	G	G	G	G	G	G	G	G	G	G
2342-4555	30	2661/3510	G	G	C	C	C	G	G	C	C	G
2342-4555	31	2704/3630	C	C	C	C	C	C	C	C	C	C
2342-4555	32	2716/3750	G	G	G	A	G	G	G	G	G	G
2342-4555	33	2846/3870	G	G	G	G	G	G	G	G	G	G
2342-4555	34	3292/3990	G	G	G	G	G	G	G	G	G	G
2342-4555	35	3470/4110	T	T	T	T	T	T	T	C	T	T
2342-4555	36	4183/4230	G	G	G	G	G	G	G	G	G	G
2342-4555	37	4201/4350	C	C	C	C	C	C	C	C	C	C
2342-4555	38	4254/4470	T	T	T	T	T	T	T	T	T	T
2342-4555	39	4384/4590	A	A	C	C	C	A	A	C	C	A
2342-4555	40	4435/4710	C	C	C	C	C	C	C	C	C	C
4651-5435	41	5180/4830	G	G	C	C	C	C	G	C	C	G
4651-5435	42	5212/4950	C	C	C	C	C	C	C	C	C	C

Table 4 (Part 3). Haplotypes of the CYP2D6 gene.

Regions Examined(a)	PS No.(b)	PS Position(c)	Haplotype Number(d)									
			21	22	23	24	25	26	27	28	29	30
378-1363	1	636/30	G	G	G	G	G	G	G	G	G	G
378-1363	2	678/150	T	T	T	T	T	T	T	T	T	T
378-1363	3	769/270	G	G	G	G	G	G	G	G	G	G
378-1363	4	776/390	A	A	A	A	A	A	A	A	A	A
378-1363	5	825/510	G	G	G	G	G	G	G	G	G	G
378-1363	6	915/630	T	T	T	T	T	T	T	T	T	T
378-1363	7	1019/750	G	G	G	G	G	G	G	G	G	G
378-1363	8	1031/870	G	G	G	G	G	G	G	G	G	G
378-1363	9	1100/990	C	C	C	C	T	T	T	T	T	T
1701-2203	10	1827/1110	G	G	G	G	G	G	G	G	G	G
1701-2203	11	1843/1230	T	T	T	T	G	G	G	G	G	G
1701-2203	12	1966/1350	G	G	G	G	G	G	G	G	G	G
1701-2203	13	1974/1470	C	C	C	C	A	A	A	A	A	A
1701-2203	14	1984/1590	A	A	A	A	A	A	A	A	G	G
1701-2203	15	1997/1710	C	C	C	C	C	G	G	G	G	G
1701-2203	16	2014/1830	T	T	T	T	T	T	T	T	T	T
1701-2203	17	2022/1950	A	A	A	A	A	A	A	A	A	A
1701-2203	18	2023/2070	C	C	C	C	C	C	C	C	C	C
1701-2203	19	2028/2190	A	A	A	A	A	A	A	A	A	A
1701-2203	20	2036/2310	T	T	T	T	T	T	T	T	T	T
1701-2203	21	2039/2430	C	C	C	C	C	C	C	C	C	C
1701-2203	22	2062/2550	A	A	A	A	A	A	A	A	A	A
1701-2203	23	2067/2670	T	T	T	T	T	T	T	T	T	T
1701-2203	24	2118/2790	C	C	C	T	C	C	C	C	C	C
1701-2203	25	2170/2910	G	G	G	G	G	G	G	G	G	G
1701-2203	26	2179/3030	G	G	G	G	G	G	G	G	G	G
2342-4555	27	2611/3150	T	T	T	T	T	T	T	T	T	T
2342-4555	28	2635/3270	T	T	T	T	T	T	T	T	T	T
2342-4555	29	2659/3390	G	G	G	G	G	G	G	G	G	G
2342-4555	30	2661/3510	C	G	G	G	C	C	C	C	C	C
2342-4555	31	2704/3630	C	C	C	C	C	C	C	C	C	C
2342-4555	32	2716/3750	G	G	G	G	G	G	G	G	G	G
2342-4555	33	2846/3870	G	G	G	G	G	A	G	G	A	G
2342-4555	34	3292/3990	G	G	G	G	G	G	G	G	G	G
2342-4555	35	3470/4110	T	T	T	T	T	T	T	T	T	T
2342-4555	36	4183/4230	G	G	G	G	G	G	G	G	G	G
2342-4555	37	4201/4350	C	C	C	C	C	C	C	C	C	C
2342-4555	38	4254/4470	T	T	T	T	T	T	T	T	T	T
2342-4555	39	4384/4590	A	A	A	A	C	C	C	C	C	C
2342-4555	40	4435/4710	C	C	C	C	C	C	A	C	C	A
4651-5435	41	5180/4830	G	C	G	C	C	C	C	C	C	C
4651-5435	42	5212/4950	C	C	C	C	C	C	C	C	C	C

Table 4 (Part 4). Haplotypes of the CYP2D6 gene.

Regions Examined(a)	PS No.(b)	PS Position(c)	Haplotype Number(d) 31	32	33	34
378-1363	1	636/30	G	G	G	G
378-1363	2	678/150	T	T	T	T
378-1363	3	769/270	G	G	G	G
378-1363	4	776/390	A	A	A	A
378-1363	5	825/510	G	G	G	G
378-1363	6	915/630	T	T	T	T
378-1363	7	1019/750	G	G	G	G
378-1363	8	1031/870	G	G	G	G
378-1363	9	1100/990	T	T	T	T
1701-2203	10	1827/1110	G	G	G	G
1701-2203	11	1843/1230	G	G	G	T
1701-2203	12	1966/1350	G	G	G	G
1701-2203	13	1974/1470	C	C	C	A
1701-2203	14	1984/1590	A	A	A	A
1701-2203	15	1997/1710	C	C	C	G
1701-2203	16	2014/1830	T	T	T	T
1701-2203	17	2022/1950	A	A	A	A
1701-2203	18	2023/2070	C	C	C	C
1701-2203	19	2028/2190	A	A	A	A
1701-2203	20	2036/2310	T	T	T	T
1701-2203	21	2039/2430	T	T	T	C
1701-2203	22	2062/2550	A	A	A	A
1701-2203	23	2067/2670	T	T	T	T
1701-2203	24	2118/2790	C	C	C	C
1701-2203	25	2170/2910	G	G	G	G
1701-2203	26	2179/3030	G	G	G	G
2342-4555	27	2611/3150	T	T	T	T
2342-4555	28	2635/3270	T	T	T	T
2342-4555	29	2659/3390	G	G	G	G
2342-4555	30	2661/3510	C	C	C	C
2342-4555	31	2704/3630	C	C	C	C
2342-4555	32	2716/3750	G	G	G	G
2342-4555	33	2846/3870	A	G	G	A
2342-4555	34	3292/3990	G	G	G	G
2342-4555	35	3470/4110	T	T	T	T
2342-4555	36	4183/4230	G	G	G	G
2342-4555	37	4201/4350	C	C	T	C
2342-4555	38	4254/4470	T	T	T	T
2342-4555	39	4384/4590	C	C	C	A
2342-4555	40	4435/4710	C	C	C	C
4651-5435	41	5180/4830	C	C	C	G
4651-5435	42	5212/4950	C	C	C	C

(a) Region examined represents the nucleotide positions defining the start and stop positions within SEQ ID NO:1 of the regions sequenced;
(b) PS = polymorphic site;
(c) Position of PS within the indicated SEQ ID NO, with the 1st position number referring to SEQ ID NO:1 and the 2nd position number referring to SEQ ID NO:151, a modified version of SEQ ID NO:1 that comprises the context sequence of each polymorphic site, PS1-PS42, to facilitate electronic searching of the haplotypes;
(d) Alleles for CYP2D6 haplotypes are presented 5′ to 3′ in each column.

SEQ ID NO:1 refers to Figure 1, with the two alternative allelic variants of each polymorphic site indicated by the appropriate nucleotide symbol. SEQ ID NO:151 is a modified version of SEQ ID NO:1 that shows the context sequence of each of PS1-PS42 in a uniform format to facilitate electronic searching of the CYP2D6 haplotypes. For each polymorphic site, SEQ ID NO:151 contains a block of 60 bases of the nucleotide sequence encompassing the centrally-located polymorphic site at the 30th position, followed by 60 bases of unspecified sequence to represent that each polymorphic site is separated by genomic sequence whose composition is defined elsewhere herein.

Table 5 below shows the percent of chromosomes characterized by a given CYP2D6 haplotype for all unrelated individuals in the Index Repository for which haplotype data was obtained. The percent of these unrelated individuals who have a given CYP2D6 haplotype pair is shown in Table 6. In Tables 5 and 6, the "Total" column shows this frequency data for all of these unrelated individuals, while the other columns show the frequency data for these unrelated individuals categorized according to their self-identified ethnogeographic origin. Abbreviations used in Tables 5 and 6 are AF = African Descent, AS = Asian, CA = Caucasian, HL = Hispanic-Latino, and AM = Native American.

Table 5. Frequency of Observed CYP2D6 Haplotypes In Unrelated Individuals

HAP No.	HAP ID	Total	CA	AF	AS	HL	AM
1	12043042	0.63	0.0	0.0	0.0	2.78	0.0
2	12043033	1.27	0.0	0.0	0.0	5.56	0.0
3	12043051	0.63	0.0	0.0	2.63	0.0	0.0
4	12043027	2.53	0.0	10.53	0.0	0.0	0.0
5	12043037	0.63	0.0	2.63	0.0	0.0	0.0
6	12043054	0.63	2.5	0.0	0.0	0.0	0.0
7	12043029	1.9	7.5	0.0	0.0	0.0	0.0
8	12043034	0.63	2.5	0.0	0.0	0.0	0.0
9	12043039	0.63	0.0	2.63	0.0	0.0	0.0
10	12043047	0.63	0.0	2.63	0.0	0.0	0.0
11	12043035	0.63	0.0	2.63	0.0	0.0	0.0
12	12043031	1.27	0.0	5.26	0.0	0.0	0.0
13	12043052	0.63	0.0	0.0	2.63	0.0	0.0
14	12043049	0.63	0.0	0.0	0.0	2.78	0.0
15	12043022	18.99	22.5	18.42	13.16	25.0	0.0
16	12043048	0.63	0.0	2.63	0.0	0.0	0.0
17	12043046	0.63	0.0	0.0	2.63	0.0	0.0
18	12043040	0.63	0.0	2.63	0.0	0.0	0.0
19	12043026	3.8	0.0	13.16	0.0	2.78	0.0
20	12043028	2.53	5.0	0.0	0.0	5.56	0.0
21	12043041	0.63	0.0	2.63	0.0	0.0	0.0
22	12043025	5.7	7.5	0.0	5.26	8.33	16.67
23	12043021	20.89	27.5	15.79	13.16	22.22	50.0
24	12043050	0.63	0.0	0.0	0.0	0.0	16.67
25	12043043	0.63	0.0	0.0	0.0	0.0	16.67
26	12043024	6.96	12.5	5.26	0.0	11.11	0.0
27	12043036	0.63	0.0	0.0	0.0	2.78	0.0
28	12043032	1.27	5.0	0.0	0.0	0.0	0.0
29	12043045	0.63	2.5	0.0	0.0	0.0	0.0
30	12043038	0.63	2.5	0.0	0.0	0.0	0.0
31	12043030	1.27	0.0	5.26	0.0	0.0	0.0
32	12043023	18.35	2.5	7.89	60.53	5.56	0.0
33	12043044	0.63	0.0	0.0	0.0	2.78	0.0
34	12043053	0.63	0.0	0.0	0.0	2.78	0.0

Table 6. Frequency of Observed CYP2D6 Haplotype Pairs In Unrelated Individuals

HAP1	HAP2	Total	CA	AF	AS	HL	AM
15	15	2.53	5.0	0.0	0.0	5.56	0.0
19	19	2.53	0.0	10.53	0.0	0.0	0.0
32	32	10.13	0.0	0.0	42.11	0.0	0.0
7	7	1.27	5.0	0.0	0.0	0.0	0.0
15	20	1.27	0.0	0.0	0.0	5.56	0.0
15	19	1.27	0.0	0.0	0.0	5.56	0.0
23	33	1.27	0.0	0.0	0.0	5.56	0.0
22	20	1.27	0.0	0.0	0.0	5.56	0.0
23	24	1.27	0.0	0.0	0.0	0.0	33.33
23	25	1.27	0.0	0.0	0.0	0.0	33.33
15	9	1.27	0.0	5.26	0.0	0.0	0.0
32	21	1.27	0.0	5.26	0.0	0.0	0.0
23	16	1.27	0.0	5.26	0.0	0.0	0.0
20	28	1.27	5.0	0.0	0.0	0.0	0.0
32	13	1.27	0.0	0.0	5.26	0.0	0.0
32	26	1.27	0.0	0.0	0.0	5.56	0.0
23	1	1.27	0.0	0.0	0.0	5.56	0.0
32	20	1.27	5.0	0.0	0.0	0.0	0.0
15	12	1.27	0.0	5.26	0.0	0.0	0.0
23	26	3.8	10.0	0.0	0.0	5.56	0.0
23	18	1.27	0.0	5.26	0.0	0.0	0.0
23	27	1.27	0.0	0.0	0.0	5.56	0.0
32	2	1.27	0.0	0.0	0.0	5.56	0.0
15	34	1.27	0.0	0.0	0.0	5.56	0.0
15	14	1.27	0.0	0.0	0.0	5.56	0.0
23	8	1.27	5.0	0.0	0.0	0.0	0.0
23	28	1.27	5.0	0.0	0.0	0.0	0.0
15	7	1.27	5.0	0.0	0.0	0.0	0.0
23	10	1.27	0.0	5.26	0.0	0.0	0.0
23	5	1.27	0.0	5.26	0.0	0.0	0.0
32	17	1.27	0.0	0.0	5.26	0.0	0.0
15	26	6.33	15.0	5.26	0.0	5.56	0.0
15	11	1.27	0.0	5.26	0.0	0.0	0.0
32	4	1.27	0.0	5.26	0.0	0.0	0.0
26	19	1.27	0.0	5.26	0.0	0.0	0.0
23	4	1.27	0.0	5.26	0.0	0.0	0.0
32	12	1.27	0.0	5.26	0.0	0.0	0.0
30	29	1.27	5.0	0.0	0.0	0.0	0.0
23	6	1.27	5.0	0.0	0.0	0.0	0.0
23	32	2.53	0.0	0.0	10.53	0.0	0.0
23	15	8.86	15.0	5.26	5.26	11.11	0.0
15	3	1.27	0.0	0.0	5.26	0.0	0.0
26	2	1.27	0.0	0.0	0.0	5.56	0.0
15	32	3.8	0.0	0.0	15.79	0.0	0.0
23	22	10.13	15.0	0.0	10.53	11.11	33.33
4	31	1.27	0.0	5.26	0.0	0.0	0.0
15	31	1.27	0.0	5.26	0.0	0.0	0.0
15	4	1.27	0.0	5.26	0.0	0.0	0.0

The size and composition of the Index Repository were chosen to represent the genetic diversity across and within four major population groups comprising the general United States population. For example, as described in Table 1 above, this repository contains approximately equal sample sizes of African-descent, Asian-American, European-American, and Hispanic-Latino population groups. Almost all individuals representing each group had all four grandparents with the same ethnogeographic background. The number of unrelated individuals in the Index Repository provides a sample size that is sufficient to detect SNPs and haplotypes that occur in the general population with high statistical certainty. For instance, a haplotype that occurs with a frequency of 5% in the general population has a probability higher than 99.9% of being observed in a sample of 80 individuals from the general population. Similarly, a haplotype that occurs with a frequency of 10% in a specific population group has a 99% probability of being observed in a sample of 20 individuals from that population group. In addition, the size and composition of the Index Repository means that the relative frequencies determined therein for the haplotypes and haplotype pairs of the CYP2D6 gene are likely to be similar to the relative frequencies of these CYP2D6 haplotypes and haplotype pairs in the general U.S. population and in the four population groups represented in the Index Repository. The genetic diversity observed for the three Native Americans is presented because it is of scientific interest, but due to the small sample size it lacks statistical significance.

In view of the above, it will be seen that the several advantages of the invention are achieved and other advantageous results attained.

As various changes could be made in the above methods and compositions without departing from the scope of the invention, it is intended that all matter contained in the above description and shown in the accompanying drawings shall be interpreted as illustrative and not in a limiting sense.

All references cited in this specification, including patents and patent applications, are hereby incorporated in their entirety by reference. The discussion of references herein is intended merely to summarize the assertions made by their authors and no admission is made that any reference constitutes prior art. Applicants reserve the right to challenge the accuracy and pertinency of the cited references.

What is Claimed is:

1. A method for haplotyping the Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) gene of an individual, which comprises determining which of the CYP2D6 haplotypes shown in the table immediately below defines one copy of the individual's CYP2D6 gene, wherein the determining step comprises identifying the phased sequence of nucleotides present at each of PS1-PS42 on at least one copy of the individual's CYP2D6 gene, and wherein each of the CYP2D6 haplotypes comprises a sequence of polymorphisms whose positions and identities are set forth in the table immediately below:

PS No.(a)	PS Position(b)	Haplotype Number(c) (Part 1)									
		1	2	3	4	5	6	7	8	9	10
1	636	A	G	G	G	G	G	G	G	G	G
2	678	T	C	T	T	T	T	T	T	T	T
3	769	G	G	C	G	G	G	G	G	G	G
4	776	A	A	G	A	A	A	A	A	A	A
5	825	G	G	G	A	G	G	G	G	G	G
6	915	T	T	T	T	C	T	T	T	T	T
7	1019	G	G	G	G	G	A	G	G	G	G
8	1031	G	G	G	G	G	G	A	A	G	G
9	1100	C	C	T	C	C	C	C	C	C	C
10	1827	G	G	G	G	G	G	G	G	C	G
11	1843	T	G	G	G	T	G	G	G	G	G
12	1966	G	G	G	G	G	G	G	G	G	A
13	1974	C	C	C	C	C	C	C	C	C	C
14	1984	A	A	A	A	A	A	A	A	A	A
15	1997	C	C	C	C	C	C	C	C	C	C
16	2014	T	T	T	T	T	T	T	T	T	T
17	2022	A	A	A	A	A	A	A	A	A	A
18	2023	C	C	C	C	C	C	C	C	T	C
19	2028	A	A	A	A	A	A	A	A	A	A
20	2036	T	T	T	T	T	T	T	T	T	T
21	2039	C	C	T	C	C	C	C	C	C	C
22	2062	A	A	A	A	A	A	A	A	A	A
23	2067	T	T	T	T	T	T	T	T	T	T
24	2118	C	C	C	C	C	C	C	C	C	C
25	2170	G	G	G	G	G	G	G	G	G	G
26	2179	G	G	G	G	G	G	G	G	G	G
27	2611	T	T	T	T	T	T	T	T	T	T
28	2635	T	T	T	T	T	T	T	T	C	T
29	2659	G	G	G	A	G	G	G	G	G	G
30	2661	G	C	C	C	G	C	C	G	C	C
31	2704	C	C	C	C	C	G	C	C	C	C
32	2716	G	G	G	G	G	G	G	G	G	G
33	2846	G	G	G	G	G	G	G	G	G	G
34	3292	G	G	G	A	G	G	G	G	G	G
35	3470	T	T	T	T	T	T	T	T	T	T
36	4183	G	G	G	A	G	G	G	G	G	G
37	4201	C	C	C	C	C	C	C	C	C	C
38	4254	T	T	T	T	T	T	T	T	T	T
39	4384	A	C	C	C	A	C	C	C	C	C
40	4435	C	C	C	C	C	C	C	C	C	C
41	5180	C	C	C	C	C	C	C	C	C	C
42	5212	C	C	C	C	C	C	C	C	C	C

PS No.(a)	PS Position(b)	Haplotype Number(c) (Part 2)									
		11	12	13	14	15	16	17	18	19	20
1	636	G	G	G	G	G	G	G	G	G	G
2	678	T	T	T	T	T	T	T	T	T	T
3	769	G	G	G	G	G	G	G	G	G	G
4	776	A	A	A	A	A	A	A	A	A	A
5	825	G	G	G	G	G	G	G	G	G	G
6	915	T	T	T	T	T	T	T	T	T	T
7	1019	G	G	G	G	G	G	G	G	G	G
8	1031	G	G	G	G	G	G	G	G	G	G
9	1100	C	C	C	C	C	C	C	C	C	C
10	1827	G	G	G	G	G	G	G	G	G	G
11	1843	G	G	G	G	G	G	G	G	G	T
12	1966	G	G	G	G	G	G	G	G	G	G
13	1974	C	C	C	C	C	C	C	C	C	C
14	1984	A	A	A	A	A	A	A	A	A	A
15	1997	C	C	C	C	C	C	C	C	C	C
16	2014	T	T	T	T	T	T	T	T	T	T
17	2022	A	A	A	A	A	A	A	A	A	A
18	2023	C	C	C	C	C	C	C	T	T	C
19	2028	A	A	A	A	A	A	A	A	A	A
20	2036	T	T	T	T	T	T	T	T	T	T
21	2039	C	C	C	C	C	C	C	C	C	C
22	2062	A	A	A	A	A	A	A	A	A	A
23	2067	G	G	T	T	T	T	T	T	T	T
24	2118	C	C	C	C	C	C	C	C	C	C
25	2170	G	G	G	G	G	G	G	G	G	A
26	2179	C	G	G	G	G	G	G	G	G	G
27	2611	T	T	A	T	T	T	T	T	T	T
28	2635	T	T	T	T	T	T	T	T	T	T
29	2659	G	G	G	G	G	G	G	G	G	G
30	2661	G	G	C	C	C	G	G	C	C	G
31	2704	C	C	C	C	C	C	C	C	C	C
32	2716	G	G	G	A	G	G	G	G	G	G
33	2846	G	G	G	G	G	G	G	G	G	G
34	3292	G	G	G	G	G	G	G	G	G	G
35	3470	T	T	T	T	T	T	T	C	T	T
36	4183	G	G	G	G	G	G	G	G	G	G
37	4201	C	C	C	C	C	C	C	C	C	C
38	4254	T	T	T	T	T	T	T	T	T	T
39	4384	A	A	C	C	C	A	A	C	C	A
40	4435	C	C	C	C	C	C	C	C	C	C
41	5180	G	G	C	C	C	C	G	C	C	G
42	5212	C	C	C	C	C	C	C	C	C	C

PS No.(a)	PS Position(b)	Haplotype Number(c) (Part 3)									
		21	22	23	24	25	26	27	28	29	30
1	636	G	G	G	G	G	G	G	G	G	G
2	678	T	T	T	T	T	T	T	T	T	T
3	769	G	G	G	G	G	G	G	G	G	G
4	776	A	A	A	A	A	A	A	A	A	A
5	825	G	G	G	G	G	G	G	G	G	G
6	915	T	T	T	T	T	T	T	T	T	T
7	1019	G	G	G	G	G	G	G	G	G	G
8	1031	G	G	G	G	G	G	G	G	G	G
9	1100	C	C	C	C	T	T	T	T	T	T
10	1827	G	G	G	G	G	G	G	G	G	G
11	1843	T	T	T	T	G	G	G	G	G	G
12	1966	G	G	G	G	G	G	G	G	G	G
13	1974	C	C	C	C	A	A	A	A	A	A
14	1984	A	A	A	A	A	A	A	A	G	G
15	1997	C	C	C	C	C	G	G	G	G	G
16	2014	T	T	T	T	T	T	T	T	T	T
17	2022	A	A	A	A	A	A	A	A	A	A
18	2023	C	C	C	C	C	C	C	C	C	C
19	2028	A	A	A	A	A	A	A	A	A	A
20	2036	T	T	T	T	T	T	T	T	T	T
21	2039	C	C	C	C	C	C	C	C	C	C
22	2062	A	A	A	A	A	A	A	A	A	A
23	2067	T	T	T	T	T	T	T	T	T	T
24	2118	C	C	C	T	C	C	C	C	C	C
25	2170	G	G	G	G	G	G	G	G	G	G
26	2179	G	G	G	G	G	G	G	G	G	G
27	2611	T	T	T	T	T	T	T	T	T	T
28	2635	T	T	T	T	T	T	T	T	T	T
29	2659	G	G	G	G	G	G	G	G	G	G
30	2661	C	G	G	G	C	C	C	C	C	C
31	2704	C	C	C	C	C	C	C	C	C	C
32	2716	G	G	G	G	G	G	G	G	G	G
33	2846	G	G	G	G	G	A	G	G	A	G
34	3292	G	G	G	G	G	G	G	G	G	G
35	3470	T	T	T	T	T	T	T	T	T	T
36	4183	G	G	G	G	G	G	G	G	G	G
37	4201	C	C	C	C	C	C	C	C	C	C
38	4254	T	T	T	T	T	T	T	T	T	T
39	4384	A	A	A	A	C	C	C	C	C	C
40	4435	C	C	C	C	C	C	A	C	C	A
41	5180	G	C	G	C	C	C	C	C	C	C
42	5212	C	C	C	C	C	C	C	C	C	C

PS No.(a)	PS Position(b)	Haplotype Number(c) (Part 4) 31	32	33	34
1	636	G	G	G	G
2	678	T	T	T	T
3	769	G	G	G	G
4	776	A	A	A	A
5	825	G	G	G	G
6	915	T	T	T	T
7	1019	G	G	G	G
8	1031	G	G	G	G
9	1100	T	T	T	T
10	1827	G	G	G	G
11	1843	G	G	G	T
12	1966	G	G	G	G
13	1974	C	C	C	A
14	1984	A	A	A	A
15	1997	C	C	C	G
16	2014	T	T	T	T
17	2022	A	A	A	A
18	2023	C	C	C	C
19	2028	A	A	A	A
20	2036	T	T	T	T
21	2039	T	T	T	C
22	2062	A	A	A	A
23	2067	T	T	T	T
24	2118	C	C	C	C
25	2170	G	G	G	G
26	2179	G	G	G	G
27	2611	T	T	T	T
28	2635	T	T	T	T
29	2659	G	G	G	G
30	2661	C	C	C	C
31	2704	C	C	C	C
32	2716	G	G	G	G
33	2846	A	G	G	A
34	3292	G	G	G	G
35	3470	T	T	T	T
36	4183	G	G	G	G
37	4201	C	C	T	C
38	4254	T	T	T	T
39	4384	C	C	C	A
40	4435	C	C	C	C
41	5180	C	C	C	G
42	5212	C	C	C	C

(a) PS = polymorphic site;
(b) Position of PS within SEQ ID NO:1;
(c) Alleles for haplotypes are presented 5′ to 3′ in each column.

2. A method for haplotyping the Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) gene of an individual, which comprises determining which of the CYP2D6 haplotype pairs

shown in the table immediately below defines both copies of the individual's CYP2D6 gene, wherein the determining step comprises identifying the phased sequence of nucleotides present at each of PS1-PS42 on both copies of the individual's CYP2D6 gene, and wherein each of the CYP2D6 haplotype pairs consists of first and second haplotypes which comprise first and second sequences of polymorphisms whose positions and identities are set forth in the table immediately below:

PS	PS	Haplotype Pair(c) (Part 1)							
No.(a)	Position(b)	15/15	19/19	32/32	7/7	15/20	15/19	23/33	22/20
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	A/A	G/G	G/G	G/G	G/G
9	1100	C/C	C/C	T/T	C/C	C/C	C/C	C/T	C/C
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	G/G	G/G	G/G	G/G	G/T	G/G	T/G	T/T
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	T/T	C/C	C/C	C/C	C/T	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/C	T/T	C/C	C/C	C/C	C/T	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	C/C	C/C	C/C	C/C	C/G	C/C	G/C	G/G
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	C/C	C/C	C/C	C/C	C/A	C/C	A/C	A/A
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	C/C	C/C	C/C	C/C	C/G	C/C	G/C	C/G
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS	PS	Haplotype Pair(c) (Part 2)							
No.(a)	Position(b)	23/24	23/25	15/9	32/21	23/16	20/28	32/13	32/26
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	C/T	C/C	T/C	C/C	C/T	T/C	T/T
10	1827	G/G	G/G	G/C	G/G	G/G	G/G	G/G	G/G
11	1843	t/t	t/g	g/g	g/t	t/g	t/g	g/g	g/g
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/A	C/C	C/C	C/C	C/A	C/C	C/A
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/C	C/C	C/C	C/G	C/C	C/G
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/C	C/C	T/C	C/C	C/C	T/C	T/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/A	T/T
28	2635	T/T	T/T	T/C	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	G/G	G/C	C/C	C/C	G/G	G/C	C/C	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	A/A	A/C	C/C	C/A	A/A	A/C	C/C	C/C
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	G/C	G/C	C/C	C/G	G/C	G/C	C/C	C/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS No.(a)	PS Position(b)	Haplotype Pair(c) (Part 3) 23/1	32/20	15/12	23/26	23/18	23/27	32/2	15/34
1	636	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	T/C	C/C	C/T	C/C	C/T	T/C	C/T
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	T/T	G/T	G/G	T/G	T/G	T/G	G/G	G/T
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	C/C	C/A	C/C	C/A	C/C	C/A
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/C	C/G	C/C	C/G	C/C	C/G
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	T/C	C/C	C/C	C/C	C/C	T/C	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/G	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	G/G	C/G	C/G	G/C	G/C	G/C	C/C	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/A
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	A/A	C/A	C/A	A/C	A/C	A/C	C/C	C/A
40	4435	C/C	C/C	C/C	C/C	C/C	C/A	C/C	C/C
41	5180	G/C	C/G	C/G	G/C	G/C	G/C	C/C	C/G
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS No.(a)	PS Position(b)	Haplotype Pair(c) (Part 4) 15/14	23/8	23/28	15/7	23/10	23/5	32/17	15/26
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
9	1100	C/C	C/C	C/T	C/C	C/C	C/C	T/C	C/T
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	G/G	T/G	T/G	G/G	T/G	T/T	G/G	G/G
12	1966	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G
13	1974	C/C	C/C	C/A	C/C	C/C	C/C	C/C	C/A
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/G	C/C	C/C	C/C	C/C	C/G
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/C	C/C	C/C	C/C	C/C	T/C	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	C/C	G/G	G/C	C/C	G/C	G/G	C/G	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	C/C	A/C	A/C	C/C	A/C	A/A	C/A	C/C
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	C/C	G/C	G/C	C/C	G/C	G/C	C/G	C/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS	PS	Haplotype Pair(c) (Part 5)							
No.(a)	Position(b)	15/11	32/4	26/19	23/4	32/12	30/29	23/6	23/32
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	T/C	T/C	C/C	T/C	T/T	C/C	C/T
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	G/G	G/G	G/G	T/G	G/G	G/G	T/G	T/G
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	A/C	C/C	C/C	A/A	C/C	C/C
14	1984	A/A	A/A	A/A	A/A	A/A	G/G	A/A	A/A
15	1997	C/C	C/C	G/C	C/C	C/C	G/G	C/C	C/C
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	T/C	C/C	C/C	T/C	C/C	C/C	C/T
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/G	T/T	T/T	T/T	T/G	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/C	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
30	2661	C/G	C/C	C/C	G/C	C/G	C/C	G/C	G/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/G	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	A/G	G/G	G/G	G/A	G/G	G/G
34	3292	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	C/A	C/C	C/C	A/C	C/A	C/C	A/C	A/C
40	4435	C/C	C/C	C/C	C/C	C/C	A/C	C/C	C/C
41	5180	C/G	C/C	C/C	G/C	C/G	C/C	G/C	G/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS	PS	Haplotype Pair(c) (Part 6)							
No.(a)	Position(b)	23/15	15/3	26/2	15/32	23/22	4/31	15/31	15/4
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/C	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/C	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/G	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	C/T	T/C	C/T	C/C	C/T	C/T	C/C
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	T/G	G/G	G/G	G/G	T/T	G/G	G/G	G/G
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	A/C	C/C	C/C	C/C	C/C	C/C
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	G/C	C/C	C/C	C/C	C/C	C/C
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/T	C/C	C/T	C/C	C/T	C/T	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
30	2661	G/C	C/C	C/C	C/C	G/G	C/C	C/C	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	A/G	G/G	G/G	G/A	G/A	G/G
34	3292	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	A/C	C/C	C/C	C/C	A/A	C/C	C/C	C/C
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	G/C	C/C	C/C	C/C	G/C	C/C	C/C	C/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

(a) PS = polymorphic site;
(b)Position of PS in SEQ ID NO:1;
(c) Haplotype pairs are represented as 1st haplotype/2nd haplotype; with alleles of each haplotype shown 5′ to 3′ as 1st polymorphism/2nd polymorphism in each column.

3. A method for genotyping the Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) gene of an individual, comprising determining for the two copies of the CYP2D6 gene present in the individual the identity of the nucleotide pair at one or more polymorphic sites (PS)

selected from the group consisting of PS1, PS2, PS3, PS4, PS6, PS10, PS12, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS22, PS23, PS24, PS26, PS27, PS28, PS29, PS32, PS34, PS36, PS37, PS38, PS39, PS40 and PS42, wherein the one or more polymorphic sites (PS) have the position and alternative alleles shown in SEQ ID NO:1.

4. The method of claim 3, wherein the determining step comprises:

(a) isolating from the individual a nucleic acid mixture comprising both copies of the CYP2D6 gene, or a fragment thereof, that are present in the individual;

(b) amplifying from the nucleic acid mixture a target region containing one of the selected polymorphic sites;

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region, wherein the oligonucleotide is designed for genotyping the selected polymorphic site in the target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized oligonucleotide in the presence of at least one terminator of the reaction, wherein the terminator is complementary to one of the alternative nucleotides present at the selected polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended oligonucleotide.

5. The method of claim 3, which comprises determining for the two copies of the CYP2D6 gene present in the individual the identity of the nucleotide pair at each of PS1-PS42.

6. A method for haplotyping the Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) gene of an individual which comprises determining, for one copy of the CYP2D6 gene present in the individual, the identity of the nucleotide at two or more polymorphic sites (PS) selected from the group consisting of PS1, PS2, PS3, PS4, PS6, PS10, PS12, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS22, PS23, PS24, PS26, PS27, PS28, PS29, PS32, PS34, PS36, PS37, PS38, PS39, PS40 and PS42, wherein the selected PS have the position and alternative alleles shown in SEQ ID NO:1.

7. The method of claim 6, further comprising determining the identity of the nucleotide at one or more polymorphic sites selected from the group consisting of PS5, PS7, PS8, PS9, PS11, PS13, PS21, PS25, PS30, PS31, PS33, PS35 and PS41, wherein the one or more polymorphic sites (PS) have the position and alternative alleles shown in SEQ ID NO:1.

8. The method of claim 6, wherein the determining step comprises:

(a) isolating from the individual a nucleic acid sample containing only one of the two copies of the CYP2D6 gene, or a fragment thereof, that is present in the individual;

(b) amplifying from the nucleic acid sample a target region containing one of the selected polymorphic sites;

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region, wherein the oligonucleotide is designed for haplotyping the selected polymorphic site in

the target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized oligonucleotide in the presence of at least one terminator of the reaction, wherein the terminator is complementary to one of the alternative nucleotides present at the selected polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended oligonucleotide.

9. A method for predicting a haplotype pair for the Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) gene of an individual comprising:

(a) identifying a CYP2D6 genotype for the individual, wherein the genotype comprises the nucleotide pair at two or more polymorphic sites (PS) selected from the group consisting of PS1, PS2, PS3, PS4, PS6, PS10, PS12, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS22, PS23, PS24, PS26, PS27, PS28, PS29, PS32, PS34, PS36, PS37, PS38, PS39, PS40 and PS42, wherein the selected PS have the position and alternative alleles shown in SEQ ID NO:1;

(b) comparing the genotype to the haplotype pair data set forth in the table immediately below; and

(c) determining which haplotype pair is consistent with the genotype of the individual and with the haplotype pair data

PS	PS	Haplotype Pair(c) (Part 1)							
No.(a)	Position(b)	15/15	19/19	32/32	7/7	15/20	15/19	23/33	22/20
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	A/A	G/G	G/G	G/G	G/G
9	1100	C/C	C/C	T/T	C/C	C/C	C/C	C/T	C/C
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	G/G	G/G	G/G	G/G	G/T	G/G	T/G	T/T
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C.	C/C	C/C	C/C	C/C	C/C	C/C	C/C
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	T/T	C/C	C/C	C/C	C/T	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/C	T/T	C/C	C/C	C/C	C/T	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	C/C	C/C	C/C	C/C	C/G	C/C	G/C	G/G
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	C/C	C/C	C/C	C/C	C/A	C/C	A/C	A/A
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	C/C	C/C	C/C	C/C	C/G	C/C	G/C	C/G
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS	PS	Haplotype Pair(c) (Part 2)							
No.(a)	Position(b)	23/24	23/25	15/9	32/21	23/16	20/28	32/13	32/26
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	C/T	C/C	T/C	C/C	C/T	T/C	T/T
10	1827	G/G	G/G	G/C	G/G	G/G	G/G	G/G	G/G
11	1843	T/T	T/G	G/G	G/T	T/G	T/G	G/G	G/G
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/A	C/C	C/C	C/C	C/A	C/C	C/A
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/C	C/C	C/C	C/G	C/C	C/G
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/C	C/C	T/C	C/C	C/C	T/C	T/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/A	T/T
28	2635	T/T	T/T	T/C	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	G/G	G/C	C/C	C/C	G/G	G/C	C/C	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	A/A	A/C	C/C	C/A	A/A	A/C	C/C	C/C
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	G/C	G/C	C/C	C/G	G/C	G/C	C/C	C/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS No.(a)	PS Position(b)	Haplotype Pair(c) (Part 3) 23/1	32/20	15/12	23/26	23/18	23/27	32/2	15/34
1	636	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	T/C	C/C	C/T	C/C	C/T	T/C	C/T
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	T/T	G/T	G/G	T/G	T/G	T/G	G/G	G/T
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	C/C	C/A	C/C	C/A	C/C	C/A
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/C	C/G	C/C	C/G	C/C	C/G
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	T/C	C/C	C/C	C/C	C/C	T/C	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/G	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	G/G	C/G	C/G	G/C	G/C	G/C	C/C	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/A
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	A/A	C/A	C/A	A/C	A/C	A/C	C/C	C/A
40	4435	C/C	C/C	C/C	C/C	C/C	C/A	C/C	C/C
41	5180	G/C	C/G	C/G	G/C	G/C	G/C	C/C	C/G
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS	PS	Haplotype Pair(c) (Part 4)							
No.(a)	Position(b)	15/14	23/8	23/28	15/7	23/10	23/5	32/17	15/26
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
9	1100	C/C	C/C	C/T	C/C	C/C	C/C	T/C	C/T
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	G/G	T/G	T/G	G/G	T/G	T/T	G/G	G/G
12	1966	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G
13	1974	C/C	C/C	C/A	C/C	C/C	C/C	C/C	C/A
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/G	C/C	C/C	C/C	C/C	C/G
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/C	C/C	C/C	C/C	C/C	T/C	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	C/C	G/G	G/C	C/C	G/C	G/G	C/G	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	C/C	A/C	A/C	C/C	A/C	A/A	C/A	C/C
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	C/C	G/C	G/C	C/C	G/C	G/C	C/G	C/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS No.(a)	PS Position(b)	Haplotype Pair(c) (Part 5) 15/11	32/4	26/19	23/4	32/12	30/29	23/6	23/32
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	T/C	T/C	C/C	T/C	T/T	C/C	C/T
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	G/G	G/G	G/G	T/G	G/G	G/G	T/G	T/G
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	A/C	C/C	C/C	A/A	C/C	C/C
14	1984	A/A	A/A	A/A	A/A	A/A	G/G	A/A	A/A
15	1997	C/C	C/C	G/C	C/C	C/C	G/G	C/C	C/C
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	T/C	C/C	C/C	T/C	C/C	C/C	C/T
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/G	T/T	T/T	T/T	T/G	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/C	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
30	2661	C/G	C/C	C/C	G/C	C/G	C/C	G/C	G/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/G	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	A/G	G/G	G/G	G/A	G/G	G/G
34	3292	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	C/A	C/C	C/C	A/C	C/A	C/C	A/C	A/C
40	4435	C/C	C/C	C/C	C/C	C/C	A/C	C/C	C/C
41	5180	C/G	C/C	C/C	G/C	C/G	C/C	G/C	G/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS No.(a)	PS Position(b)	Haplotype Pair(c) (Part 6) 23/15	15/3	26/2	15/32	23/22	4/31	15/31	15/4
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/C	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/C	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/G	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	C/T	T/C	C/T	C/C	C/T	C/T	C/C
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	T/G	G/G	G/G	G/G	T/T	G/G	G/G	G/G
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	A/C	C/C	C/C	C/C	C/C	C/C
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	G/C	C/C	C/C	C/C	C/C	C/C
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/T	C/C	C/T	C/C	C/T	C/T	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
30	2661	G/C	C/C	C/C	C/C	G/G	C/C	C/C	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	A/G	G/G	G/G	G/A	G/A	G/G
34	3292	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	A/C	C/C	C/C	C/C	A/A	C/C	C/C	C/C
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	G/C	C/C	C/C	C/C	G/C	C/C	C/C	C/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

(a) PS = polymorphic site;
(b) Position of PS in SEQ ID NO:1;
(c) Haplotype pairs are represented as 1st haplotype/2nd haplotype; with alleles of each haplotype shown 5′ to 3′ as 1st polymorphism/2nd polymorphism in each column.

10. The method of claim 9, wherein the identified genotype of the individual comprises the nucleotide pair at each of PS1-PS42, which have the position and alternative alleles shown in SEQ ID NO:1.

11. A method for identifying an association between a trait and at least one haplotype or haplotype pair of the Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) gene which comprises comparing the frequency of the haplotype or haplotype pair in a population exhibiting the trait with the frequency of the haplotype or haplotype pair in a reference population, wherein the haplotype is selected from haplotypes 1-34 shown in the table presented immediately below, wherein each of the haplotypes comprises a sequence of polymorphisms whose positions and identities are set forth in the table immediately below:

PS No.(a)	PS Position(b)	Haplotype Number(c) (Part 1)									
		1	2	3	4	5	6	7	8	9	10
1	636	A	G	G	G	G	G	G	G	G	G
2	678	T	C	T	T	T	T	T	T	T	T
3	769	G	G	C	G	G	G	G	G	G	G
4	776	A	A	G	A	A	A	A	A	A	A
5	825	G	G	G	A	G	G	G	G	G	G
6	915	T	T	T	T	C	T	T	T	T	T
7	1019	G	G	G	G	G	A	G	G	G	G
8	1031	G	G	G	G	G	G	A	A	G	G
9	1100	C	C	T	C	C	C	C	C	C	C
10	1827	G	G	G	G	G	G	G	G	C	G
11	1843	T	G	G	G	T	G	G	G	G	G
12	1966	G	G	G	G	G	G	G	G	G	A
13	1974	C	C	C	C	C	C	C	C	C	C
14	1984	A	A	A	A	A	A	A	A	A	A
15	1997	C	C	C	C	C	C	C	C	C	C
16	2014	T	T	T	T	T	T	T	T	T	T
17	2022	A	A	A	A	A	A	A	A	A	A
18	2023	C	C	C	C	C	C	C	C	T	C
19	2028	A	A	A	A	A	A	A	A	A	A
20	2036	T	T	T	T	T	T	T	T	T	T
21	2039	C	C	T	C	C	C	C	C	C	C
22	2062	A	A	A	A	A	A	A	A	A	A
23	2067	T	T	T	T	T	T	T	T	T	T
24	2118	C	C	C	C	C	C	C	C	C	C
25	2170	G	G	G	G	G	G	G	G	G	G
26	2179	G	G	G	G	G	G	G	G	G	G
27	2611	T	T	T	T	T	T	T	T	T	T
28	2635	T	T	T	T	T	T	T	T	C	T
29	2659	G	G	G	A	G	G	G	G	G	G
30	2661	G	C	C	C	G	C	C	G	C	C
31	2704	C	C	C	C	C	G	C	C	C	C
32	2716	G	G	G	G	G	G	G	G	G	G
33	2846	G	G	G	G	G	G	G	G	G	G
34	3292	G	G	G	A	G	G	G	G	G	G
35	3470	T	T	T	T	T	T	T	T	T	T
36	4183	G	G	G	A	G	G	G	G	G	G
37	4201	C	C	C	C	C	C	C	C	C	C
38	4254	T	T	T	T	T	T	T	T	T	T
39	4384	A	C	C	C	A	C	C	C	C	C
40	4435	C	C	C	C	C	C	C	C	C	C
41	5180	C	C	C	C	C	C	C	C	C	C
42	5212	C	C	C	C	C	C	C	C	C	C

PS	PS	Haplotype Number(c) (Part 2)									
No.(a)	Position(b)	11	12	13	14	15	16	17	18	19	20
1	636	G	G	G	G	G	G	G	G	G	G
2	678	T	T	T	T	T	T	T	T	T	T
3	769	G	G	G	G	G	G	G	G	G	G
4	776	A	A	A	A	A	A	A	A	A	A
5	825	G	G	G	G	G	G	G	G	G	G
6	915	T	T	T	T	T	T	T	T	T	T
7	1019	G	G	G	G	G	G	G	G	G	G
8	1031	G	G	G	G	G	G	G	G	G	G
9	1100	C	C	C	C	C	C	C	C	C	C
10	1827	G	G	G	G	G	G	G	G	G	G
11	1843	G	G	G	G	G	G	G	G	G	T
12	1966	G	G	G	G	G	G	G	G	G	G
13	1974	C	C	C	C	C	C	C	C	C	C
14	1984	A	A	A	A	A	A	A	A	A	A
15	1997	C	C	C	C	C	C	C	C	C	C
16	2014	T	T	T	T	T	T	T	T	T	T
17	2022	A	A	A	A	A	A	A	A	A	A
18	2023	C	C	C	C	C	C	C	T	T	C
19	2028	A	A	A	A	A	A	A	A	A	A
20	2036	T	T	T	T	T	T	T	T	T	T
21	2039	C	C	C	C	C	C	C	C	C	C
22	2062	A	A	A	A	A	A	A	A	A	A
23	2067	G	G	T	T	T	T	T	T	T	T
24	2118	C	C	C	C	C	C	C	C	C	C
25	2170	G	G	G	G	G	G	G	G	G	A
26	2179	C	G	G	G	G	G	G	G	G	G
27	2611	T	T	A	T	T	T	T	T	T	T
28	2635	T	T	T	T	T	T	T	T	T	T
29	2659	G	G	G	G	G	G	G	G	G	G
30	2661	G	G	C	C	C	G	G	C	C	G
31	2704	C	C	C	C	C	C	C	C	C	C
32	2716	G	G	G	A	G	G	G	G	G	G
33	2846	G	G	G	G	G	G	G	G	G	G
34	3292	G	G	G	G	G	G	G	G	G	G
35	3470	T	T	T	T	T	T	T	C	T	T
36	4183	G	G	G	G	G	G	G	G	G	G
37	4201	C	C	C	C	C	C	C	C	C	C
38	4254	T	T	T	T	T	T	T	T	T	T
39	4384	A	A	C	C	C	A	A	C	C	A
40	4435	C	C	C	C	C	C	C	C	C	C
41	5180	G	G	C	C	C	C	G	C	C	G
42	5212	C	C	C	C	C	C	C	C	C	C

PS No.(a)	PS Position(b)	Haplotype Number(c) (Part 3)									
		21	22	23	24	25	26	27	28	29	30
1	636	G	G	G	G	G	G	G	G	G	G
2	678	T	T	T	T	T	T	T	T	T	T
3	769	G	G	G	G	G	G	G	G	G	G
4	776	A	A	A	A	A	A	A	A	A	A
5	825	G	G	G	G	G	G	G	G	G	G
6	915	T	T	T	T	T	T	T	T	T	T
7	1019	G	G	G	G	G	G	G	G	G	G
8	1031	G	G	G	G	G	G	G	G	G	G
9	1100	C	C	C	C	T	T	T	T	T	T
10	1827	G	G	G	G	G	G	G	G	G	G
11	1843	T	T	T	T	G	G	G	G	G	G
12	1966	G	G	G	G	G	G	G	G	G	G
13	1974	C	C	C	C	A	A	A	A	A	A
14	1984	A	A	A	A	A	A	A	A	G	G
15	1997	C	C	C	C	C	G	G	G	G	G
16	2014	T	T	T	T	T	T	T	T	T	T
17	2022	A	A	A	A	A	A	A	A	A	A
18	2023	C	C	C	C	C	C	C	C	C	C
19	2028	A	A	A	A	A	A	A	A	A	A
20	2036	T	T	T	T	T	T	T	T	T	T
21	2039	C	C	C	C	C	C	C	C	C	C
22	2062	A	A	A	A	A	A	A	A	A	A
23	2067	T	T	T	T	T	T	T	T	T	T
24	2118	C	C	C	T	C	C	C	C	C	C
25	2170	G	G	G	G	G	G	G	G	G	G
26	2179	G	G	G	G	G	G	G	G	G	G
27	2611	T	T	T	T	T	T	T	T	T	T
28	2635	T	T	T	T	T	T	T	T	T	T
29	2659	G	G	G	G	G	G	G	G	G	G
30	2661	C	G	G	G	C	C	C	C	C	C
31	2704	C	C	C	C	C	C	C	C	C	C
32	2716	G	G	G	G	G	G	G	G	G	G
33	2846	G	G	G	G	G	A	G	G	A	G
34	3292	G	G	G	G	G	G	G	G	G	G
35	3470	T	T	T	T	T	T	T	T	T	T
36	4183	G	G	G	G	G	G	G	G	G	G
37	4201	C	C	C	C	C	C	C	C	C	C
38	4254	T	T	T	T	T	T	T	T	T	T
39	4384	A	A	A	A	C	C	C	C	C	C
40	4435	C	C	C	C	C	C	A	C	C	A
41	5180	G	C	G	C	C	C	C	C	C	C
42	5212	C	C	C	C	C	C	C	C	C	C

PS No.(a)	PS Position(b)	Haplotype Number(c) (Part 4) 31	32	33	34
1	636	G	G	G	G
2	678	T	T	T	T
3	769	G	G	G	G
4	776	A	A	A	A
5	825	G	G	G	G
6	915	T	T	T	T
7	1019	G	G	G	G
8	1031	G	G	G	G
9	1100	T	T	T	T
10	1827	G	G	G	G
11	1843	G	G	G	T
12	1966	G	G	G	G
13	1974	C	C	C	A
14	1984	A	A	A	A
15	1997	C	C	C	G
16	2014	T	T	T	T
17	2022	A	A	A	A
18	2023	C	C	C	C
19	2028	A	A	A	A
20	2036	T	T	T	T
21	2039	T	T	T	C
22	2062	A	A	A	A
23	2067	T	T	T	T
24	2118	C	C	C	C
25	2170	G	G	G	G
26	2179	G	G	G	G
27	2611	T	T	T	T
28	2635	T	T	T	T
29	2659	G	G	G	G
30	2661	C	C	C	C
31	2704	C	C	C	C
32	2716	G	G	G	G
33	2846	A	G	G	A
34	3292	G	G	G	G
35	3470	T	T	T	T
36	4183	G	G	G	G
37	4201	C	C	T	C
38	4254	T	T	T	T
39	4384	C	C	C	A
40	4435	C	C	C	C
41	5180	C	C	C	G
42	5212	C	C	C	C

(a) PS = polymorphic site;
(b) Position of PS within SEQ ID NO:1;
(c) Alleles for haplotypes are presented 5′ to 3′ in each column;

and wherein the haplotype pair is selected from the haplotype pairs shown in the table immediately below, wherein each of the CYP2D6 haplotype pairs consists of first and second haplotypes which comprise first and second sequences of polymorphisms whose positions in

SEQ ID NO:1 and identities are set forth in the table immediately below:

PS No.(a)	PS Position(b)	Haplotype Pair(c) (Part 1) 15/15	19/19	32/32	7/7	15/20	15/19	23/33	22/20
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	A/A	G/G	G/G	G/G	G/G
9	1100	C/C	C/C	T/T	C/C	C/C	C/C	C/T	C/C
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	G/G	G/G	G/G	G/G	G/T	G/G	T/G	T/T
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	T/T	C/C	C/C	C/C	C/T	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/C	T/T	C/C	C/C	C/C	C/T	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	C/C	C/C	C/C	C/C	C/G	C/C	G/C	G/G
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	C/C	C/C	C/C	C/C	C/A	C/C	A/C	A/A
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	C/C	C/C	C/C	C/C	C/G	C/C	G/C	C/G
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS No.(a)	PS Position(b)	Haplotype Pair(c) (Part 2) 23/24	23/25	15/9	32/21	23/16	20/28	32/13	32/26
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	C/T	C/C	T/C	C/C	C/T	T/C	T/T
10	1827	G/G	G/G	G/C	G/G	G/G	G/G	G/G	G/G
11	1843	T/T	T/G	G/G	G/T	T/G	T/G	G/G	G/G
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/A	C/C	C/C	C/C	C/A	C/C	C/A
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/C	C/C	C/C	C/G	C/C	C/G
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/C	C/C	T/C	C/C	C/C	T/C	T/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/A	T/T
28	2635	T/T	T/T	T/C	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	G/G	G/C	C/C	C/C	G/G	G/C	C/C	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	A/A	A/C	C/C	C/A	A/A	A/C	C/C	C/C
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	G/C	G/C	C/C	C/G	G/C	G/C	C/C	C/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS No.(a)	PS Position(b)	Haplotype Pair(c) (Part 3)							
		23/1	32/20	15/12	23/26	23/18	23/27	32/2	15/34
1	636	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	T/C	C/C	C/T	C/C	C/T	T/C	C/T
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	T/T	G/T	G/G	T/G	T/G	T/G	G/G	G/T
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	C/C	C/A	C/C	C/A	C/C	C/A
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/C	C/G	C/C	C/G	C/C	C/G
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	T/C	C/C	C/C	C/C	C/C	T/C	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/G	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	G/G	C/G	C/G	G/C	G/C	G/C	C/C	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/A
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	A/A	C/A	C/A	A/C	A/C	A/C	C/C	C/A
40	4435	C/C	C/C	C/C	C/C	C/C	C/A	C/C	C/C
41	5180	G/C	C/G	C/G	G/C	G/C	G/C	C/C	C/G
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS No.(a)	PS Position(b)	Haplotype Pair(c) (Part 4) 15/14	23/8	23/28	15/7	23/10	23/5	32/17	15/26
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
9	1100	C/C	C/C	C/T	C/C	C/C	C/C	T/C	C/T
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	G/G	T/G	T/G	G/G	T/G	T/T	G/G	G/G
12	1966	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G
13	1974	C/C	C/C	C/A	C/C	C/C	C/C	C/C	C/A
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/G	C/C	C/C	C/C	C/C	C/G
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/C	C/C	C/C	C/C	C/C	T/C	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	C/C	G/G	G/C	C/C	G/C	G/G	C/G	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	C/C	A/C	A/C	C/C	A/C	A/A	C/A	C/C
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	C/C	G/C	G/C	C/C	G/C	G/C	C/G	C/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS No.(a)	PS Position(b)	Haplotype Pair(c) (Part 5) 15/11	32/4	26/19	23/4	32/12	30/29	23/6	23/32
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	T/C	T/C	C/C	T/C	T/T	C/C	C/T
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	G/G	G/G	G/G	T/G	G/G	G/G	T/G	T/G
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	A/C	C/C	C/C	A/A	C/C	C/C
14	1984	A/A	A/A	A/A	A/A	A/A	G/G	A/A	A/A
15	1997	C/C	C/C	G/C	C/C	C/C	G/G	C/C	C/C
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	T/C	C/C	C/C	T/C	C/C	C/C	C/T
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/G	T/T	T/T	T/T	T/G	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/C	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
30	2661	C/G	C/C	C/C	G/C	C/G	C/C	G/C	G/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/G	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	A/G	G/G	G/G	G/A	G/G	G/G
34	3292	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	C/A	C/C	C/C	A/C	C/A	C/C	A/C	A/C
40	4435	C/C	C/C	C/C	C/C	C/C	A/C	C/C	C/C
41	5180	C/G	C/C	C/C	G/C	C/G	C/C	G/C	G/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS	PS	Haplotype Pair(c) (Part 6)							
No.(a)	Position(b)	23/15	15/3	26/2	15/32	23/22	4/31	15/31	15/4
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/C	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/C	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/G	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	C/T	T/C	C/T	C/C	C/T	C/T	C/C
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	T/G	G/G	G/G	G/G	T/T	G/G	G/G	G/G
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	A/C	C/C	C/C	C/C	C/C	C/C
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	G/C	C/C	C/C	C/C	C/C	C/C
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/T	C/C	C/T	C/C	C/T	C/T	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
30	2661	G/C	C/C	C/C	C/C	G/G	C/C	C/C	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	A/G	G/G	G/G	G/A	G/A	G/G
34	3292	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	A/C	C/C	C/C	C/C	A/A	C/C	C/C	C/C
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	G/C	C/C	C/C	C/C	G/C	C/C	C/C	C/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

(a) PS = polymorphic site;
(b) Position of PS in SEQ ID NO:1;
(c) Haplotype pairs are represented as 1st haplotype/2nd haplotype; with alleles of each haplotype shown 5′ to 3′ as 1st polymorphism/2nd polymorphism in each column;

wherein a higher frequency of the haplotype or haplotype pair in the trait population than in the reference population indicates the trait is associated with the haplotype or haplotype pair.

12. The method of claim 11, wherein the trait is a clinical response to a drug targeting or metabolized by CYP2D6 or to a drug for treating a condition or disease associated with CYP2D6 activity.

13. An isolated oligonucleotide designed for detecting a polymorphism in the Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) gene at a polymorphic site (PS) selected from the group consisting of PS1, PS2, PS3, PS4, PS6, PS10, PS12, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS22, PS23, PS24, PS26, PS27, PS28, PS29, PS32, PS34, PS36, PS37, PS38, PS39, PS40 and PS42, wherein the selected PS have the position and alternative alleles shown in SEQ ID NO:1.

14. The isolated oligonucleotide of claim 13, which is an allele-specific oligonucleotide that specifically hybridizes to an allele of the CYP2D6 gene at a region containing the polymorphic site.

15. The allele-specific oligonucleotide of claim 14, which comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS:4-32, the complements of SEQ ID NOS:4-32, and SEQ ID NOS:33-90.

16. The isolated oligonucleotide of claim 13, which is a primer-extension oligonucleotide.

17. The primer-extension oligonucleotide of claim 16,which comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS:91-148.

18. A kit for haplotyping or genotyping the Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) gene of an individual, which comprises a set of oligonucleotides designed to haplotype or genotype each of polymorphic sites (PS) PS1, PS2, PS3, PS4, PS6, PS10, PS12, PS14, PS15, PS16, PS17, PS18, PS19, PS20, PS22, PS23, PS24, PS26, PS27, PS28, PS29, PS32, PS34, PS36, PS37, PS38, PS39, PS40 and PS42, wherein the selected PS have the position and alternative alleles shown in SEQ ID NO:1.

19. The kit of claim 18, which further comprises oligonucleotides designed to genotype or haplotype each of PS5, PS7, PS8, PS9, PS11, PS13, PS21, PS25, PS30, PS31, PS33, PS35 and PS41, wherein the selected PS have the position and alternative alleles shown in SEQ ID NO:1.

20. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) a first nucleotide sequence which comprises a Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) isogene, wherein the CYP2D6 isogene is selected from the group consisting of isogenes 1-34 shown in the table immediately below and wherein each of the isogenes comprises the regions of SEQ ID NO:1 shown in the table immediately below and wherein each of the isogenes 1-34 is further defined by the corresponding sequence of polymorphisms whose positions and identities are set forth in the table immediately below; and

Region Examined(a)	PS No.(b)	PS Position(c)	Isogene Number(d) (Part 1) 1	2	3	4	5	6	7	8	9	10
378-1363	1	636	A	G	G	G	G	G	G	G	G	G
378-1363	2	678	T	C	T	T	T	T	T	T	T	T
378-1363	3	769	G	G	C	G	G	G	G	G	G	G
378-1363	4	776	A	A	G	A	A	A	A	A	A	A
378-1363	5	825	G	G	G	A	G	G	G	G	G	G
378-1363	6	915	T	T	T	T	C	T	T	T	T	T
378-1363	7	1019	G	G	G	G	G	A	G	G	G	G
378-1363	8	1031	G	G	G	G	G	G	A	A	G	G
378-1363	9	1100	C	C	T	C	C	C	C	C	C	C
1701-2203	10	1827	G	G	G	G	G	G	G	G	C	G
1701-2203	11	1843	T	G	G	G	T	G	G	G	G	G
1701-2203	12	1966	G	G	G	G	G	G	G	G	G	A
1701-2203	13	1974	C	C	C	C	C	C	C	C	C	C
1701-2203	14	1984	A	A	A	A	A	A	A	A	A	A
1701-2203	15	1997	C	C	C	C	C	C	C	C	C	C
1701-2203	16	2014	T	T	T	T	T	T	T	T	T	T
1701-2203	17	2022	A	A	A	A	A	A	A	A	A	A
1701-2203	18	2023	C	C	C	C	C	C	C	C	T	C
1701-2203	19	2028	A	A	A	A	A	A	A	A	A	A
1701-2203	20	2036	T	T	T	T	T	T	T	T	T	T
1701-2203	21	2039	C	C	T	C	C	C	C	C	C	C
1701-2203	22	2062	A	A	A	A	A	A	A	A	A	A
1701-2203	23	2067	T	T	T	T	T	T	T	T	T	T
1701-2203	24	2118	C	C	C	C	C	C	C	C	C	C
1701-2203	25	2170	G	G	G	G	G	G	G	G	G	G
1701-2203	26	2179	G	G	G	G	G	G	G	G	G	G
2342-4555	27	2611	T	T	T	T	T	T	T	T	T	T
2342-4555	28	2635	T	T	T	T	T	T	T	T	C	T
2342-4555	29	2659	G	G	G	A	G	G	G	G	G	G
2342-4555	30	2661	G	C	C	C	G	C	C	G	C	C
2342-4555	31	2704	C	C	C	C	C	G	C	C	C	C
2342-4555	32	2716	G	G	G	G	G	G	G	G	G	G
2342-4555	33	2846	G	G	G	G	G	G	G	G	G	G
2342-4555	34	3292	G	G	G	A	G	G	G	G	G	G
2342-4555	35	3470	T	T	T	T	T	T	T	T	T	T
2342-4555	36	4183	G	G	G	A	G	G	G	G	G	G
2342-4555	37	4201	C	C	C	C	C	C	C	C	C	C
2342-4555	38	4254	T	T	T	T	T	T	T	T	T	T
2342-4555	39	4384	A	C	C	C	A	C	C	C	C	C
2342-4555	40	4435	C	C	C	C	C	C	C	C	C	C
4651-5435	41	5180	C	C	C	C	C	C	C	C	C	C
4651-5435	42	5212	C	C	C	C	C	C	C	C	C	C

Region Examined(a)	PS No.(b)	PS Position(c)	Isogene Number(d) (Part 2) 11	12	13	14	15	16	17	18	19	20
378-1363	1	636	G	G	G	G	G	G	G	G	G	G
378-1363	2	678	T	T	T	T	T	T	T	T	T	T
378-1363	3	769	G	G	G	G	G	G	G	G	G	G
378-1363	4	776	A	A	A	A	A	A	A	A	A	A
378-1363	5	825	G	G	G	G	G	G	G	G	G	G
378-1363	6	915	T	T	T	T	T	T	T	T	T	T
378-1363	7	1019	G	G	G	G	G	G	G	G	G	G
378-1363	8	1031	G	G	G	G	G	G	G	G	G	G
378-1363	9	1100	C	C	C	C	C	C	C	C	C	C
1701-2203	10	1827	G	G	G	G	G	G	G	G	G	G
1701-2203	11	1843	G	G	G	G	G	G	G	G	G	T
1701-2203	12	1966	G	G	G	G	G	G	G	G	G	G
1701-2203	13	1974	C	C	C	C	C	C	C	C	C	C
1701-2203	14	1984	A	A	A	A	A	A	A	A	A	A
1701-2203	15	1997	C	C	C	C	C	C	C	C	C	C
1701-2203	16	2014	T	T	T	T	T	T	T	T	T	T
1701-2203	17	2022	A	A	A	A	A	A	A	A	A	A
1701-2203	18	2023	C	C	C	C	C	C	C	T	T	C
1701-2203	19	2028	A	A	A	A	A	A	A	A	A	A
1701-2203	20	2036	T	T	T	T	T	T	T	T	T	T
1701-2203	21	2039	C	C	C	C	C	C	C	C	C	C
1701-2203	22	2062	A	A	A	A	A	A	A	A	A	A
1701-2203	23	2067	G	G	T	T	T	T	T	T	T	T
1701-2203	24	2118	C	C	C	C	C	C	C	C	C	C
1701-2203	25	2170	G	G	G	G	G	G	G	G	G	A
1701-2203	26	2179	C	G	G	G	G	G	G	G	G	G
2342-4555	27	2611	T	T	A	T	T	T	T	T	T	T
2342-4555	28	2635	T	T	T	T	T	T	T	T	T	T
2342-4555	29	2659	G	G	G	G	G	G	G	G	G	G
2342-4555	30	2661	G	G	C	C	C	G	G	C	C	G
2342-4555	31	2704	C	C	C	C	C	C	C	C	C	C
2342-4555	32	2716	G	G	G	A	G	G	G	G	G	G
2342-4555	33	2846	G	G	G	G	G	G	G	G	G	G
2342-4555	34	3292	G	G	G	G	G	G	G	G	G	G
2342-4555	35	3470	T	T	T	T	T	T	T	C	T	T
2342-4555	36	4183	G	G	G	G	G	G	G	G	G	G
2342-4555	37	4201	C	C	C	C	C	C	C	C	C	C
2342-4555	38	4254	T	T	T	T	T	T	T	T	T	T
2342-4555	39	4384	A	A	C	C	C	A	A	C	C	A
2342-4555	40	4435	C	C	C	C	C	C	C	C	C	C
4651-5435	41	5180	G	G	C	C	C	C	G	C	C	G
4651-5435	42	5212	C	C	C	C	C	C	C	C	C	C

Region Examined(a)	PS No.(b)	PS Position(c)	Isogene Number(d) (Part 3) 21	22	23	24	25	26	27	28	29	30
378-1363	1	636	G	G	G	G	G	G	G	G	G	G
378-1363	2	678	T	T	T	T	T	T	T	T	T	T
378-1363	3	769	G	G	G	G	G	G	G	G	G	G
378-1363	4	776	A	A	A	A	A	A	A	A	A	A
378-1363	5	825	G	G	G	G	G	G	G	G	G	G
378-1363	6	915	T	T	T	T	T	T	T	T	T	T
378-1363	7	1019	G	G	G	G	G	G	G	G	G	G
378-1363	8	1031	G	G	G	G	G	G	G	G	G	G
378-1363	9	1100	C	C	C	C	T	T	T	T	T	T
1701-2203	10	1827	G	G	G	G	G	G	G	G	G	G
1701-2203	11	1843	T	T	T	T	G	G	G	G	G	G
1701-2203	12	1966	G	G	G	G	G	G	G	G	G	G
1701-2203	13	1974	C	C	C	C	A	A	A	A	A	A
1701-2203	14	1984	A	A	A	A	A	A	A	A	G	G
1701-2203	15	1997	C	C	C	C	C	G	G	G	G	G
1701-2203	16	2014	T	T	T	T	T	T	T	T	T	T
1701-2203	17	2022	A	A	A	A	A	A	A	A	A	A
1701-2203	18	2023	C	C	C	C	C	C	C	C	C	C
1701-2203	19	2028	A	A	A	A	A	A	A	A	A	A
1701-2203	20	2036	T	T	T	T	T	T	T	T	T	T
1701-2203	21	2039	C	C	C	C	C	C	C	C	C	C
1701-2203	22	2062	A	A	A	A	A	A	A	A	A	A
1701-2203	23	2067	T	T	T	T	T	T	T	T	T	T
1701-2203	24	2118	C	C	C	T	C	C	C	C	C	C
1701-2203	25	2170	G	G	G	G	G	G	G	G	G	G
1701-2203	26	2179	G	G	G	G	G	G	G	G	G	G
2342-4555	27	2611	T	T	T	T	T	T	T	T	T	T
2342-4555	28	2635	T	T	T	T	T	T	T	T	T	T
2342-4555	29	2659	G	G	G	G	G	G	G	G	G	G
2342-4555	30	2661	C	G	G	G	C	C	C	C	C	C
2342-4555	31	2704	C	C	C	C	C	C	C	C	C	C
2342-4555	32	2716	G	G	G	G	G	G	G	G	G	G
2342-4555	33	2846	G	G	G	G	G	A	G	G	A	G
2342-4555	34	3292	G	G	G	G	G	G	G	G	G	G
2342-4555	35	3470	T	T	T	T	T	T	T	T	T	T
2342-4555	36	4183	G	G	G	G	G	G	G	G	G	G
2342-4555	37	4201	C	C	C	C	C	C	C	C	C	C
2342-4555	38	4254	T	T	T	T	T	T	T	T	T	T
2342-4555	39	4384	A	A	A	A	C	C	C	C	C	C
2342-4555	40	4435	C	C	C	C	C	C	A	C	C	A
4651-5435	41	5180	G	C	G	C	C	C	C	C	C	C
4651-5435	42	5212	C	C	C	C	C	C	C	C	C	C

Region Examined(a)	PS No.(b)	PS Position(c)	Isogene Number(d) (Part 4) 31	32	33	34
378-1363	1	636	G	G	G	G
378-1363	2	678	T	T	T	T
378-1363	3	769	G	G	G	G
378-1363	4	776	A	A	A	A
378-1363	5	825	G	G	G	G
378-1363	6	915	T	T	T	T
378-1363	7	1019	G	G	G	G
378-1363	8	1031	G	G	G	G
378-1363	9	1100	T	T	T	T
1701-2203	10	1827	G	G	G	G
1701-2203	11	1843	G	G	G	T
1701-2203	12	1966	G	G	G	G
1701-2203	13	1974	C	C	C	A
1701-2203	14	1984	A	A	A	A
1701-2203	15	1997	C	C	C	G
1701-2203	16	2014	T	T	T	T
1701-2203	17	2022	A	A	A	A
1701-2203	18	2023	C	C	C	C
1701-2203	19	2028	A	A	A	A
1701-2203	20	2036	T	T	T	T
1701-2203	21	2039	T	T	T	C
1701-2203	22	2062	A	A	A	A
1701-2203	23	2067	T	T	T	T
1701-2203	24	2118	C	C	C	C
1701-2203	25	2170	G	G	G	G
1701-2203	26	2179	G	G	G	G
2342-4555	27	2611	T	T	T	T
2342-4555	28	2635	T	T	T	T
2342-4555	29	2659	G	G	G	G
2342-4555	30	2661	C	C	C	C
2342-4555	31	2704	C	C	C	C
2342-4555	32	2716	G	G	G	G
2342-4555	33	2846	A	G	G	A
2342-4555	34	3292	G	G	G	G
2342-4555	35	3470	T	T	T	T
2342-4555	36	4183	G	G	G	G
2342-4555	37	4201	C	C	T	C
2342-4555	38	4254	T	T	T	T
2342-4555	39	4384	C	C	C	A
2342-4555	40	4435	C	C	C	C
4651-5435	41	5180	C	C	C	G
4651-5435	42	5212	C	C	C	C

(a) Alleles for isogenes are presented 5′ to 3′ in each column;
(b) PS = polymorphic site;
(c) Position of PS in SEQ ID NO:1;
(d) Region examined represents the nucleotide positions defining the start and stop positions within the 1st SEQ ID NO of the sequenced region.

(b) a second nucleotide sequence which is complementary to the first nucleotide sequence.

21. The isolated polynucleotide of claim 20, which is a DNA molecule and comprises both the first

and second nucleotide sequences and further comprises expression regulatory elements operably linked to the first nucleotide sequence.

22. A recombinant nonhuman organism transformed or transfected with the isolated polynucleotide of claim 21, wherein the organism expresses a CYP2D6 protein that is encoded by the first nucleotide sequence.

23. The recombinant nonhuman organism of claim 22, which is a transgenic animal.

24. An isolated fragment of a Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) isogene, wherein the fragment comprises at least 10 nucleotides in one of the regions of SEQ ID NO:1 shown in the table immediately below and wherein the fragment comprises one or more polymorphisms selected from the group consisting of adenine at PS1, cytosine at PS2, cytosine at PS3, guanine at PS4, cytosine at PS6, cytosine at PS10, adenine at PS12, guanine at PS14, guanine at PS15, cytosine at PS16, thymine at PS17, thymine at PS18, guanine at PS19, cytosine at PS20, guanine at PS22, guanine at PS23, thymine at PS24, cytosine at PS26, adenine at PS27, cytosine at PS28, adenine at PS29, adenine at PS32, adenine at PS34, adenine at PS36, thymine at PS37, cytosine at PS38, cytosine at PS39, adenine at PS40 and thymine at PS42, wherein the selected polymorphism has the position set forth in the table immediately below:

Region Examined(a)	PS No.(b)	PS Position(c)	Isogene Number(d) (Part 1) 1	2	3	4	5	6	7	8	9	10
378-1363	1	636	A	G	G	G	G	G	G	G	G	G
378-1363	2	678	T	C	T	T	T	T	T	T	T	T
378-1363	3	769	G	G	C	G	G	G	G	G	G	G
378-1363	4	776	A	A	G	A	A	A	A	A	A	A
378-1363	5	825	G	G	G	A	G	G	G	G	G	G
378-1363	6	915	T	T	T	T	C	T	T	T	T	T
378-1363	7	1019	G	G	G	G	G	A	G	G	G	G
378-1363	8	1031	G	G	G	G	G	G	A	A	G	G
378-1363	9	1100	C	C	T	C	C	C	C	C	C	C
1701-2203	10	1827	G	G	G	G	G	G	G	G	C	G
1701-2203	11	1843	T	G	G	G	T	G	G	G	G	G
1701-2203	12	1966	G	G	G	G	G	G	G	G	G	A
1701-2203	13	1974	C	C	C	C	C	C	C	C	C	C
1701-2203	14	1984	A	A	A	A	A	A	A	A	A	A
1701-2203	15	1997	C	C	C	C	C	C	C	C	C	C
1701-2203	16	2014	T	T	T	T	T	T	T	T	T	T
1701-2203	17	2022	A	A	A	A	A	A	A	A	A	A
1701-2203	18	2023	C	C	C	C	C	C	C	C	T	C
1701-2203	19	2028	A	A	A	A	A	A	A	A	A	A
1701-2203	20	2036	T	T	T	T	T	T	T	T	T	T
1701-2203	21	2039	C	C	T	C	C	C	C	C	C	C
1701-2203	22	2062	A	A	A	A	A	A	A	A	A	A
1701-2203	23	2067	T	T	T	T	T	T	T	T	T	T
1701-2203	24	2118	C	C	C	C	C	C	C	C	C	C
1701-2203	25	2170	G	G	G	G	G	G	G	G	G	G
1701-2203	26	2179	G	G	G	G	G	G	G	G	G	G
2342-4555	27	2611	T	T	T	T	T	T	T	T	T	T
2342-4555	28	2635	T	T	T	T	T	T	T	T	C	T
2342-4555	29	2659	G	G	G	A	G	G	G	G	G	G
2342-4555	30	2661	G	C	C	C	G	C	C	G	C	C
2342-4555	31	2704	C	C	C	C	C	G	C	C	C	C
2342-4555	32	2716	G	G	G	G	G	G	G	G	G	G
2342-4555	33	2846	G	G	G	G	G	G	G	G	G	G
2342-4555	34	3292	G	G	G	A	G	G	G	G	G	G
2342-4555	35	3470	T	T	T	T	T	T	T	T	T	T
2342-4555	36	4183	G	G	G	A	G	G	G	G	G	G
2342-4555	37	4201	C	C	C	C	C	C	C	C	C	C
2342-4555	38	4254	T	T	T	T	T	T	T	T	T	T
2342-4555	39	4384	A	C	C	C	A	C	C	C	C	C
2342-4555	40	4435	C	C	C	C	C	C	C	C	C	C
4651-5435	41	5180	C	C	C	C	C	C	C	C	C	C
4651-5435	42	5212	C	C	C	C	C	C	C	C	C	C

Region Examined(a)	PS No.(b)	PS Position(c)	Isogene Number(d) (Part 2) 11	12	13	14	15	16	17	18	19	20
378-1363	1	636	G	G	G	G	G	G	G	G	G	G
378-1363	2	678	T	T	T	T	T	T	T	T	T	T
378-1363	3	769	G	G	G	G	G	G	G	G	G	G
378-1363	4	776	A	A	A	A	A	A	A	A	A	A
378-1363	5	825	G	G	G	G	G	G	G	G	G	G
378-1363	6	915	T	T	T	T	T	T	T	T	T	T
378-1363	7	1019	G	G	G	G	G	G	G	G	G	G
378-1363	8	1031	G	G	G	G	G	G	G	G	G	G
378-1363	9	1100	C	C	C	C	C	C	C	C	C	C
1701-2203	10	1827	G	G	G	G	G	G	G	G	G	G
1701-2203	11	1843	G	G	G	G	G	G	G	G	G	T
1701-2203	12	1966	G	G	G	G	G	G	G	G	G	G
1701-2203	13	1974	C	C	C	C	C	C	C	C	C	C
1701-2203	14	1984	A	A	A	A	A	A	A	A	A	A
1701-2203	15	1997	C	C	C	C	C	C	C	C	C	C
1701-2203	16	2014	T	T	T	T	T	T	T	T	T	T
1701-2203	17	2022	A	A	A	A	A	A	A	A	A	A
1701-2203	18	2023	C	C	C	C	C	C	C	T	T	C
1701-2203	19	2028	A	A	A	A	A	A	A	A	A	A
1701-2203	20	2036	T	T	T	T	T	T	T	T	T	T
1701-2203	21	2039	C	C	C	C	C	C	C	C	C	C
1701-2203	22	2062	A	A	A	A	A	A	A	A	A	A
1701-2203	23	2067	G	G	T	T	T	T	T	T	T	T
1701-2203	24	2118	C	C	C	C	C	C	C	C	C	C
1701-2203	25	2170	G	G	G	G	G	G	G	G	G	A
1701-2203	26	2179	C	G	G	G	G	G	G	G	G	G
2342-4555	27	2611	T	T	A	T	T	T	T	T	T	T
2342-4555	28	2635	T	T	T	T	T	T	T	T	T	T
2342-4555	29	2659	G	G	G	G	G	G	G	G	G	G
2342-4555	30	2661	G	G	C	C	C	G	G	C	C	G
2342-4555	31	2704	C	C	C	C	C	C	C	C	C	C
2342-4555	32	2716	G	G	G	A	G	G	G	G	G	G
2342-4555	33	2846	G	G	G	G	G	G	G	G	G	G
2342-4555	34	3292	G	G	G	G	G	G	G	G	G	G
2342-4555	35	3470	T	T	T	T	T	T	T	C	T	T
2342-4555	36	4183	G	G	G	G	G	G	G	G	G	G
2342-4555	37	4201	C	C	C	C	C	C	C	C	C	C
2342-4555	38	4254	T	T	T	T	T	T	T	T	T	T
2342-4555	39	4384	A	A	C	C	C	A	A	C	C	A
2342-4555	40	4435	C	C	C	C	C	C	C	C	C	C
4651-5435	41	5180	G	G	C	C	C	C	G	C	C	G
4651-5435	42	5212	C	C	C	C	C	C	C	C	C	C

Region Examined(a)	PS No.(b)	PS Position(c)	Isogene Number(d) (Part 3) 21	22	23	24	25	26	27	28	29	30
378-1363	1	636	G	G	G	G	G	G	G	G	G	G
378-1363	2	678	T	T	T	T	T	T	T	T	T	T
378-1363	3	769	G	G	G	G	G	G	G	G	G	G
378-1363	4	776	A	A	A	A	A	A	A	A	A	A
378-1363	5	825	G	G	G	G	G	G	G	G	G	G
378-1363	6	915	T	T	T	T	T	T	T	T	T	T
378-1363	7	1019	G	G	G	G	G	G	G	G	G	G
378-1363	8	1031	G	G	G	G	G	G	G	G	G	G
378-1363	9	1100	C	C	C	C	T	T	T	T	T	T
1701-2203	10	1827	G	G	G	G	G	G	G	G	G	G
1701-2203	11	1843	T	T	T	T	G	G	G	G	G	G
1701-2203	12	1966	G	G	G	G	G	G	G	G	G	G
1701-2203	13	1974	C	C	C	C	A	A	A	A	A	A
1701-2203	14	1984	A	A	A	A	A	A	A	A	G	G
1701-2203	15	1997	C	C	C	C	C	G	G	G	G	G
1701-2203	16	2014	T	T	T	T	T	T	T	T	T	T
1701-2203	17	2022	A	A	A	A	A	A	A	A	A	A
1701-2203	18	2023	C	C	C	C	C	C	C	C	C	C
1701-2203	19	2028	A	A	A	A	A	A	A	A	A	A
1701-2203	20	2036	T	T	T	T	T	T	T	T	T	T
1701-2203	21	2039	C	C	C	C	C	C	C	C	C	C
1701-2203	22	2062	A	A	A	A	A	A	A	A	A	A
1701-2203	23	2067	T	T	T	T	T	T	T	T	T	T
1701-2203	24	2118	C	C	C	T	C	C	C	C	C	C
1701-2203	25	2170	G	G	G	G	G	G	G	G	G	G
1701-2203	26	2179	G	G	G	G	G	G	G	G	G	G
2342-4555	27	2611	T	T	T	T	T	T	T	T	T	T
2342-4555	28	2635	T	T	T	T	T	T	T	T	T	T
2342-4555	29	2659	G	G	G	G	G	G	G	G	G	G
2342-4555	30	2661	C	G	G	G	C	C	C	C	C	C
2342-4555	31	2704	C	C	C	C	C	C	C	C	C	C
2342-4555	32	2716	G	G	G	G	G	G	G	G	G	G
2342-4555	33	2846	G	G	G	G	G	A	G	G	A	G
2342-4555	34	3292	G	G	G	G	G	G	G	G	G	G
2342-4555	35	3470	T	T	T	T	T	T	T	T	T	T
2342-4555	36	4183	G	G	G	G	G	G	G	G	G	G
2342-4555	37	4201	C	C	C	C	C	C	C	C	C	C
2342-4555	38	4254	T	T	T	T	T	T	T	T	T	T
2342-4555	39	4384	A	A	A	A	C	C	C	C	C	C
2342-4555	40	4435	C	C	C	C	C	C	A	C	C	A
4651-5435	41	5180	G	C	G	C	C	C	C	C	C	C
4651-5435	42	5212	C	C	C	C	C	C	C	C	C	C

Region Examined(a)	PS No.(b)	PS Position(c)	Isogene Number(d) (Part 4) 31	32	33	34
378-1363	1	636	G	G	G	G
378-1363	2	678	T	T	T	T
378-1363	3	769	G	G	G	G
378-1363	4	776	A	A	A	A
378-1363	5	825	G	G	G	G
378-1363	6	915	T	T	T	T
378-1363	7	1019	G	G	G	G
378-1363	8	1031	G	G	G	G
378-1363	9	1100	T	T	T	T
1701-2203	10	1827	G	G	G	G
1701-2203	11	1843	G	G	G	T
1701-2203	12	1966	G	G	G	G
1701-2203	13	1974	C	C	C	A
1701-2203	14	1984	A	A	A	A
1701-2203	15	1997	C	C	C	G
1701-2203	16	2014	T	T	T	T
1701-2203	17	2022	A	A	A	A
1701-2203	18	2023	C	C	C	C
1701-2203	19	2028	A	A	A	A
1701-2203	20	2036	T	T	T	T
1701-2203	21	2039	T	T	T	C
1701-2203	22	2062	A	A	A	A
1701-2203	23	2067	T	T	T	T
1701-2203	24	2118	C	C	C	C
1701-2203	25	2170	G	G	G	G
1701-2203	26	2179	G	G	G	G
2342-4555	27	2611	T	T	T	T
2342-4555	28	2635	T	T	T	T
2342-4555	29	2659	G	G	G	G
2342-4555	30	2661	C	C	C	C
2342-4555	31	2704	C	C	C	C
2342-4555	32	2716	G	G	G	G
2342-4555	33	2846	A	G	G	A
2342-4555	34	3292	G	G	G	G
2342-4555	35	3470	T	T	T	T
2342-4555	36	4183	G	G	G	G
2342-4555	37	4201	C	C	T	C
2342-4555	38	4254	T	T	T	T
2342-4555	39	4384	C	C	C	A
2342-4555	40	4435	C	C	C	C
4651-5435	41	5180	C	C	C	G
4651-5435	42	5212	C	C	C	C

(a) Region examined represents the nucleotide positions defining the start and stop positions within SEQ ID NO:1 of the regions sequenced;
(b) PS = polymorphic site;
(c) Position of PS within SEQ ID NO:1;
(d) Alleles for CYP2D6 isogenes are presented 5′ to 3′ in each column.

25. An isolated polynucleotide comprising a coding sequence of a CYP2D6 isogene, wherein the coding sequence comprises the regions of SEQ ID NO:2 that are defined by exons 1-9, except at

each of the polymorphic sites which have the positions in SEQ ID NO:2 and polymorphisms set forth in the table immediately below:

PS	PS	Isogene Coding Sequence Number(c) (Part 1)									
No.(a)	Position(b)	1c	2c	3c	4c	5c	6c	7c	8c	9c	10c
7	19	G	G	G	G	G	A	G	G	G	G
8	31	G	G	G	G	G	G	A	A	G	G
9	100	C	C	T	C	C	C	C	C	C	C
12	263	G	G	G	G	G	G	G	G	G	A
13	271	C	C	C	C	C	C	C	C	C	C
14	281	A	A	A	A	A	A	A	A	A	A
15	294	C	C	C	C	C	C	C	C	C	C
16	311	T	T	T	T	T	T	T	T	T	T
17	319	A	A	A	A	A	A	A	A	A	A
18	320	C	C	C	C	C	C	C	C	T	C
19	325	A	A	A	A	A	A	A	A	A	A
20	333	T	T	T	T	T	T	T	T	T	T
21	336	C	C	T	C	C	C	C	C	C	C
27	358	T	T	T	T	T	T	T	T	T	T
28	382	T	T	T	T	T	T	T	T	C	T
29	406	G	G	G	A	G	G	G	G	G	G
30	408	G	C	C	C	G	C	C	G	C	C
31	451	C	C	C	C	C	G	C	C	C	C
32	463	G	G	G	G	G	G	G	G	G	G
35	696	T	T	T	T	T	T	T	T	T	T
36	1012	G	G	G	A	G	G	G	G	G	G
37	1030	C	C	C	C	C	C	C	C	C	C
38	1083	T	T	T	T	T	T	T	T	T	T
41	1457	C	C	C	C	C	C	C	C	C	C
42	1489	C	C	C	C	C	C	C	C	C	C

PS No.(a)	PS Position(b)	Isogene Coding Sequence Number(c) (Part 2) 13c	14c	15c	16c	18c	19c	21c	22c	24c	25c
7	19	G	G	G	G	G	G	G	G	G	G
8	31	G	G	G	G	G	G	G	G	G	G
9	100	C	C	C	C	C	C	C	C	C	T
12	263	G	G	G	G	G	G	G	G	G	G
13	271	C	C	C	C	C	C	C	C	C	A
14	281	A	A	A	A	A	A	A	A	A	A
15	294	C	C	C	C	C	C	C	C	C	C
16	311	T	T	T	T	T	T	T	T	T	T
17	319	A	A	A	A	A	A	A	A	A	A
18	320	C	C	C	C	T	T	C	C	C	C
19	325	A	A	A	A	A	A	A	A	A	A
20	333	T	T	T	T	T	T	T	T	T	T
21	336	C	C	C	C	C	C	C	C	C	C
27	358	A	T	T	T	T	T	T	T	T	T
28	382	T	T	T	T	T	T	T	T	T	T
29	406	G	G	G	G	G	G	G	G	G	G
30	408	C	C	C	G	C	C	C	G	G	C
31	451	C	C	C	C	C	C	C	C	C	C
32	463	G	A	G	G	G	G	G	G	G	G
35	696	T	T	T	T	C	T	T	T	T	T
36	1012	G	G	G	G	G	G	G	G	G	G
37	1030	C	C	C	C	C	C	C	C	C	C
38	1083	T	T	T	T	T	T	T	T	T	T
41	1457	C	C	C	C	C	C	G	C	C	C
42	1489	C	C	C	C	C	C	C	C	C	C

PS No.(a)	PS Position(b)	Isogene Coding Sequence Number(c) (Part 3) 26c	27c	28c	29c	30c	31c	32c	33c	34c
7	19	G	G	G	G	G	G	G	G	G
8	31	G	G	G	G	G	G	G	G	G
9	100	T	T	T	T	T	T	T	T	T
12	263	G	G	G	G	G	G	G	G	G
13	271	A	A	A	A	A	C	C	C	A
14	281	A	A	A	G	G	A	A	A	A
15	294	G	G	G	G	G	C	C	C	G
16	311	T	T	T	T	T	T	T	T	T
17	319	A	A	A	A	A	A	A	A	A
18	320	C	C	C	C	C	C	C	C	C
19	325	A	A	A	A	A	A	A	A	A
20	333	T	T	T	T	T	T	T	T	T
21	336	C	C	C	C	C	T	T	T	C
27	358	T	T	T	T	T	T	T	T	T
28	382	T	T	T	T	T	T	T	T	T
29	406	G	G	G	G	G	G	G	G	G
30	408	C	C	C	C	C	C	C	C	C
31	451	C	C	C	C	C	C	C	C	C
32	463	G	G	G	G	G	G	G	G	G
35	696	T	T	T	T	T	T	T	T	T
36	1012	G	G	G	G	G	G	G	G	G
37	1030	C	C	C	C	C	C	C	T	C
38	1083	T	T	T	T	T	T	T	T	T
41	1457	C	C	C	C	C	C	C	C	G
42	1489	C	C	C	C	C	C	C	C	C

(a) PS = polymorphic site;
(b) Position of PS in SEQ ID NO:2;
(c) Alleles for the isogene coding sequence are presented 5′ to 3′ in each column; the numerical portion of the isogene coding sequence number represents the number of the parent full CYP2D6 isogene.

26. A recombinant nonhuman organism transformed or transfected with the isolated polynucleotide of claim 25, wherein the organism expresses a Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) protein that is encoded by the polymorphic variant sequence.

27. The recombinant nonhuman organism of claim 26, which is a transgenic animal.

28. An isolated fragment of a CYP2D6 coding sequence, wherein the fragment comprises one or more polymorphisms selected from the group consisting of adenine at a position corresponding to nucleotide 263, guanine at a position corresponding to nucleotide 281, guanine at a position corresponding to nucleotide 294, cytosine at a position corresponding to nucleotide 311, thymine at a position corresponding to nucleotide 319, thymine at a position corresponding to nucleotide 320, guanine at a position corresponding to nucleotide 325, cytosine at a position corresponding to nucleotide 333, adenine at a position corresponding to nucleotide 358, cytosine at a position corresponding to nucleotide 382, adenine at a position corresponding to nucleotide 406, adenine at a position corresponding to nucleotide 463, adenine at a position corresponding to nucleotide 1012, thymine at a position corresponding to nucleotide 1030, cytosine at a position corresponding to nucleotide 1083 and thymine at a position corresponding to nucleotide 1489 in SEQ ID NO:2.

29 An isolated polypeptide comprising an amino acid sequence which is a polymorphic variant of a reference sequence for the Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) protein, wherein the reference sequence comprises SEQ ID NO:3 for the regions encoded by exons 1-9, except the polymorphic variant comprises one or more variant amino acids selected from the group consisting of histidine at a position corresponding to amino acid position 88, arginine at a position corresponding to amino acid position 94, alanine at a position corresponding to amino acid position 104, phenylalanine at a position corresponding to amino acid position 107, phenylalanine at a position corresponding to amino acid position 107, valine at a position corresponding to amino acid position 109, isoleucine at a position corresponding to amino acid position 120, arginine at a position corresponding to amino acid position 128, isoleucine at a position corresponding to amino acid position 136, lysine at a position corresponding to amino acid position 155, methionine at a position corresponding to amino acid position 338, termination codon at a position corresponding to amino acid position 344 and cysteine at a position corresponding to amino acid position 497.

30. An isolated monoclonal antibody specific for and immunoreactive with the isolated polypeptide

of claim 29.

31. A method for screening for drugs, or other chemical compounds, that bind to or are enzymatic substrates for the isolated polypeptide of claim 29 which comprises contacting the CYP2D6 polymorphic variant with a candidate agent and assaying for binding activity.

32. An isolated fragment of a CYP2D6 protein, wherein the fragment comprises one or more variant amino acids selected from the group consisting of histidine at a position corresponding to amino acid position 88, arginine at a position corresponding to amino acid position 94, alanine at a position corresponding to amino acid position 104, phenylalanine at a position corresponding to amino acid position 107, phenylalanine at a position corresponding to amino acid position 107, valine at a position corresponding to amino acid position 109, isoleucine at a position corresponding to amino acid position 120, arginine at a position corresponding to amino acid position 128, isoleucine at a position corresponding to amino acid position 136, lysine at a position corresponding to amino acid position 155, methionine at a position corresponding to amino acid position 338, termination codon at a position corresponding to amino acid position 344 and cysteine at a position corresponding to amino acid position 497 in SEQ ID NO:3.

33. A computer system for storing and analyzing polymorphism data for the Cytochrome P450, subfamily IID, Polypeptide 6 gene, comprising:

(a) a central processing unit (CPU);

(b) a communication interface;

(c) a display device;

(d) an input device; and

(e) a database containing the polymorphism data;

wherein the polymorphism data comprises any one or more of the haplotypes set forth in the table immediately below:

PS No.(a)	PS Position(b)	Haplotype Number(c) (Part 1) 1	2	3	4	5	6	7	8	9	10
1	636	A	G	G	G	G	G	G	G	G	G
2	678	T	C	T	T	T	T	T	T	T	T
3	769	G	G	C	G	G	G	G	G	G	G
4	776	A	A	G	A	A	A	A	A	A	A
5	825	G	G	G	A	G	G	G	G	G	G
6	915	T	T	T	T	C	T	T	T	T	T
7	1019	G	G	G	G	G	A	G	G	G	G
8	1031	G	G	G	G	G	G	A	A	G	G
9	1100	C	C	T	C	C	C	C	C	C	C
10	1827	G	G	G	G	G	G	G	G	C	G
11	1843	T	G	G	G	T	G	G	G	G	G
12	1966	G	G	G	G	G	G	G	G	G	A
13	1974	C	C	C	C	C	C	C	C	C	C
14	1984	A	A	A	A	A	A	A	A	A	A
15	1997	C	C	C	C	C	C	C	C	C	C
16	2014	T	T	T	T	T	T	T	T	T	T
17	2022	A	A	A	A	A	A	A	A	A	A
18	2023	C	C	C	C	C	C	C	C	T	C
19	2028	A	A	A	A	A	A	A	A	A	A
20	2036	T	T	T	T	T	T	T	T	T	T
21	2039	C	C	T	C	C	C	C	C	C	C
22	2062	A	A	A	A	A	A	A	A	A	A
23	2067	T	T	T	T	T	T	T	T	T	T
24	2118	C	C	C	C	C	C	C	C	C	C
25	2170	G	G	G	G	G	G	G	G	G	G
26	2179	G	G	G	G	G	G	G	G	G	G
27	2611	T	T	T	T	T	T	T	T	T	T
28	2635	T	T	T	T	T	T	T	T	C	T
29	2659	G	G	G	A	G	G	G	G	G	G
30	2661	G	C	C	C	G	C	C	G	C	C
31	2704	C	C	C	C	C	G	C	C	C	C
32	2716	G	G	G	G	G	G	G	G	G	G
33	2846	G	G	G	G	G	G	G	G	G	G
34	3292	G	G	G	A	G	G	G	G	G	G
35	3470	T	T	T	T	T	T	T	T	T	T
36	4183	G	G	G	A	G	G	G	G	G	G
37	4201	C	C	C	C	C	C	C	C	C	C
38	4254	T	T	T	T	T	T	T	T	T	T
39	4384	A	C	C	C	A	C	C	C	C	C
40	4435	C	C	C	C	C	C	C	C	C	C
41	5180	C	C	C	C	C	C	C	C	C	C
42	5212	C	C	C	C	C	C	C	C	C	C

PS	PS	Haplotype Number(c) (Part 2)									
No.(a)	Position(b)	11	12	13	14	15	16	17	18	19	20
1	636	G	G	G	G	G	G	G	G	G	G
2	678	T	T	T	T	T	T	T	T	T	T
3	769	G	G	G	G	G	G	G	G	G	G
4	776	A	A	A	A	A	A	A	A	A	A
5	825	G	G	G	G	G	G	G	G	G	G
6	915	T	T	T	T	T	T	T	T	T	T
7	1019	G	G	G	G	G	G	G	G	G	G
8	1031	G	G	G	G	G	G	G	G	G	G
9	1100	C	C	C	C	C	C	C	C	C	C
10	1827	G	G	G	G	G	G	G	G	G	G
11	1843	G	G	G	G	G	G	G	G	G	T
12	1966	G	G	G	G	G	G	G	G	G	G
13	1974	C	C	C	C	C	C	C	C	C	C
14	1984	A	A	A	A	A	A	A	A	A	A
15	1997	C	C	C	C	C	C	C	C	C	C
16	2014	T	T	T	T	T	T	T	T	T	T
17	2022	A	A	A	A	A	A	A	A	A	A
18	2023	C	C	C	C	C	C	C	T	T	C
19	2028	A	A	A	A	A	A	A	A	A	A
20	2036	T	T	T	T	T	T	T	T	T	T
21	2039	C	C	C	C	C	C	C	C	C	C
22	2062	A	A	A	A	A	A	A	A	A	A
23	2067	G	G	T	T	T	T	T	T	T	T
24	2118	C	C	C	C	C	C	C	C	C	C
25	2170	G	G	G	G	G	G	G	G	G	A
26	2179	C	G	G	G	G	G	G	G	G	G
27	2611	T	T	A	T	T	T	T	T	T	T
28	2635	T	T	T	T	T	T	T	T	T	T
29	2659	G	G	G	G	G	G	G	G	G	G
30	2661	G	G	C	C	C	G	G	C	C	G
31	2704	C	C	C	C	C	C	C	C	C	C
32	2716	G	G	G	A	G	G	G	G	G	G
33	2846	G	G	G	G	G	G	G	G	G	G
34	3292	G	G	G	G	G	G	G	G	G	G
35	3470	T	T	T	T	T	T	T	C	T	T
36	4183	G	G	G	G	G	G	G	G	G	G
37	4201	C	C	C	C	C	C	C	C	C	C
38	4254	T	T	T	T	T	T	T	T	T	T
39	4384	A	A	C	C	C	A	A	C	C	A
40	4435	C	C	C	C	C	C	C	C	C	C
41	5180	G	G	C	C	C	C	G	C	C	G
42	5212	C	C	C	C	C	C	C	C	C	C

PS No.(a)	PS Position(b)	Haplotype Number(c) (Part 3) 21	22	23	24	25	26	27	28	29	30
1	636	G	G	G	G	G	G	G	G	G	G
2	678	T	T	T	T	T	T	T	T	T	T
3	769	G	G	G	G	G	G	G	G	G	G
4	776	A	A	A	A	A	A	A	A	A	A
5	825	G	G	G	G	G	G	G	G	G	G
6	915	T	T	T	T	T	T	T	T	T	T
7	1019	G	G	G	G	G	G	G	G	G	G
8	1031	G	G	G	G	G	G	G	G	G	G
9	1100	C	C	C	C	T	T	T	T	T	T
10	1827	G	G	G	G	G	G	G	G	G	G
11	1843	T	T	T	T	G	G	G	G	G	G
12	1966	G	G	G	G	G	G	G	G	G	G
13	1974	C	C	C	C	A	A	A	A	A	A
14	1984	A	A	A	A	A	A	A	A	G	G
15	1997	C	C	C	C	C	G	G	G	G	G
16	2014	T	T	T	T	T	T	T	T	T	T
17	2022	A	A	A	A	A	A	A	A	A	A
18	2023	C	C	C	C	C	C	C	C	C	C
19	2028	A	A	A	A	A	A	A	A	A	A
20	2036	T	T	T	T	T	T	T	T	T	T
21	2039	C	C	C	C	C	C	C	C	C	C
22	2062	A	A	A	A	A	A	A	A	A	A
23	2067	T	T	T	T	T	T	T	T	T	T
24	2118	C	C	C	T	C	C	C	C	C	C
25	2170	G	G	G	G	G	G	G	G	G	G
26	2179	G	G	G	G	G	G	G	G	G	G
27	2611	T	T	T	T	T	T	T	T	T	T
28	2635	T	T	T	T	T	T	T	T	T	T
29	2659	G	G	G	G	G	G	G	G	G	G
30	2661	C	G	G	G	C	C	C	C	C	C
31	2704	C	C	C	C	C	C	C	C	C	C
32	2716	G	G	G	G	G	G	G	G	G	G
33	2846	G	G	G	G	G	A	G	G	A	G
34	3292	G	G	G	G	G	G	G	G	G	G
35	3470	T	T	T	T	T	T	T	T	T	T
36	4183	G	G	G	G	G	G	G	G	G	G
37	4201	C	C	C	C	C	C	C	C	C	C
38	4254	T	T	T	T	T	T	T	T	T	T
39	4384	A	A	A	A	C	C	C	C	C	C
40	4435	C	C	C	C	C	C	A	C	C	A
41	5180	G	C	G	C	C	C	C	C	C	C
42	5212	C	C	C	C	C	C	C	C	C	C

PS No.(a)	PS Position(b)	Haplotype Number(c) (Part 4) 31	32	33	34
1	636	G	G	G	G
2	678	T	T	T	T
3	769	G	G	G	G
4	776	A	A	A	A
5	825	G	G	G	G
6	915	T	T	T	T
7	1019	G	G	G	G
8	1031	G	G	G	G
9	1100	T	T	T	T
10	1827	G	G	G	G
11	1843	G	G	G	T
12	1966	G	G	G	G
13	1974	C	C	C	A
14	1984	A	A	A	A
15	1997	C	C	C	G
16	2014	T	T	T	T
17	2022	A	A	A	A
18	2023	C	C	C	C
19	2028	A	A	A	A
20	2036	T	T	T	T
21	2039	T	T	T	C
22	2062	A	A	A	A
23	2067	T	T	T	T
24	2118	C	C	C	C
25	2170	G	G	G	G
26	2179	G	G	G	G
27	2611	T	T	T	T
28	2635	T	T	T	T
29	2659	G	G	G	G
30	2661	C	C	C	C
31	2704	C	C	C	C
32	2716	G	G	G	G
33	2846	A	G	G	A
34	3292	G	G	G	G
35	3470	T	T	T	T
36	4183	G	G	G	G
37	4201	C	C	T	C
38	4254	T	T	T	T
39	4384	C	C	C	A
40	4435	C	C	C	C
41	5180	C	C	C	G
42	5212	C	C	C	C

(a) PS = polymorphic site;
(b) Position of PS within SEQ ID NO:1;
(c) Alleles for haplotypes are presented 5′ to 3′ in each column;

the haplotype pairs set forth in the table immediately below:

PS	PS	Haplotype Pair(c) (Part 1)							
No.(a)	Position(b)	15/15	19/19	32/32	7/7	15/20	15/19	23/33	22/20
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	A/A	G/G	G/G	G/G	G/G
9	1100	C/C	C/C	T/T	C/C	C/C	C/C	C/T	C/C
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	G/G	G/G	G/G	G/G	G/T	G/G	T/G	T/T
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	T/T	C/C	C/C	C/C	C/T	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/C	T/T	C/C	C/C	C/C	C/T	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	C/C	C/C	C/C	C/C	C/G	C/C	G/C	G/G
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	C/C	C/C	C/C	C/C	C/A	C/C	A/C	A/A
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	C/C	C/C	C/C	C/C	C/G	C/C	G/C	C/G
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS	PS	Haplotype Pair(c) (Part 2)							
No.(a)	Position(b)	23/24	23/25	15/9	32/21	23/16	20/28	32/13	32/26
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	C/T	C/C	T/C	C/C	C/T	T/C	T/T
10	1827	G/G	G/G	G/C	G/G	G/G	G/G	G/G	G/G
11	1843	T/T	T/G	G/G	G/T	T/G	T/G	G/G	G/G
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/A	C/C	C/C	C/C	C/A	C/C	C/A
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/C	C/C	C/C	C/G	C/C	C/G
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/C	C/C	T/C	C/C	C/C	T/C	T/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/A	T/T
28	2635	T/T	T/T	T/C	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	G/G	G/C	C/C	C/C	G/G	G/C	C/C	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	A/A	A/C	C/C	C/A	A/A	A/C	C/C	C/C
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	G/C	G/C	C/C	C/G	G/C	G/C	C/C	C/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS	PS	Haplotype Pair(c) (Part 3)							
No.(a)	Position(b)	23/1	32/20	15/12	23/26	23/18	23/27	32/2	15/34
1	636	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	T/C	C/C	C/T	C/C	C/T	T/C	C/T
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	T/T	G/T	G/G	T/G	T/G	T/G	G/G	G/T
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	C/C	C/A	C/C	C/A	C/C	C/A
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/C	C/G	C/C	C/G	C/C	C/G
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	T/C	C/C	C/C	C/C	C/C	T/C	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/G	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	G/G	C/G	C/G	G/C	G/C	G/C	C/C	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/A
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	A/A	C/A	C/A	A/C	A/C	A/C	C/C	C/A
40	4435	C/C	C/C	C/C	C/C	C/C	C/A	C/C	C/C
41	5180	G/C	C/G	C/G	G/C	G/C	G/C	C/C	C/G
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS	PS	Haplotype Pair(c) (Part 4)							
No.(a)	Position(b)	15/14	23/8	23/28	15/7	23/10	23/5	32/17	15/26
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
6	915	T/T	T/T,	T/T	T/T	T/T	T/C	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
9	1100	C/C	C/C	C/T	C/C	C/C	C/C	T/C	C/T
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	G/G	T/G	T/G	G/G	T/G	T/T	G/G	G/G
12	1966	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G
13	1974	C/C	C/C	C/A	C/C	C/C	C/C	C/C	C/A
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	C/G	C/C	C/C	C/C	C/C	C/G
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/C	C/C	C/C	C/C	C/C	T/C	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
30	2661	C/C	G/G	G/C	C/C	G/C	G/G	C/G	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A
34	3292	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	C/C	A/C	A/C	C/C	A/C	A/A	C/A	C/C
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	C/C	G/C	G/C	C/C	G/C	G/C	C/G	C/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS	PS	Haplotype Pair(c) (Part 5)							
No.(a)	Position(b)	15/11	32/4	26/19	23/4	32/12	30/29	23/6	23/32
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	T/C	T/C	C/C	T/C	T/T	C/C	C/T
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	G/G	G/G	G/G	T/G	G/G	G/G	T/G	T/G
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	A/C	C/C	C/C	A/A	C/C	C/C
14	1984	A/A	A/A	A/A	A/A	A/A	G/G	A/A	A/A
15	1997	C/C	C/C	G/C	C/C	C/C	G/G	C/C	C/C
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	T/C	C/C	C/C	T/C	C/C	C/C	C/T
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/G	T/T	T/T	T/T	T/G	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/C	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
30	2661	C/G	C/C	C/C	G/C	C/G	C/C	G/C	G/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/G	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	A/G	G/G	G/G	G/A	G/G	G/G
34	3292	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	C/A	C/C	C/C	A/C	C/A	C/C	A/C	A/C
40	4435	C/C	C/C	C/C	C/C	C/C	A/C	C/C	C/C
41	5180	C/G	C/C	C/C	G/C	C/G	C/C	G/C	G/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

PS	PS	Haplotype Pair(c) (Part 6)							
No.(a)	Position(b)	23/15	15/3	26/2	15/32	23/22	4/31	15/31	15/4
1	636	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
2	678	T/T	T/T	T/C	T/T	T/T	T/T	T/T	T/T
3	769	G/G	G/C	G/G	G/G	G/G	G/G	G/G	G/G
4	776	A/A	A/G	A/A	A/A	A/A	A/A	A/A	A/A
5	825	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
6	915	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
7	1019	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
8	1031	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
9	1100	C/C	C/T	T/C	C/T	C/C	C/T	C/T	C/C
10	1827	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
11	1843	T/G	G/G	G/G	G/G	T/T	G/G	G/G	G/G
12	1966	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
13	1974	C/C	C/C	A/C	C/C	C/C	C/C	C/C	C/C
14	1984	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
15	1997	C/C	C/C	G/C	C/C	C/C	C/C	C/C	C/C
16	2014	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
17	2022	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
18	2023	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
19	2028	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
20	2036	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
21	2039	C/C	C/T	C/C	C/T	C/C	C/T	C/T	C/C
22	2062	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
23	2067	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
24	2118	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
25	2170	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
26	2179	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
27	2611	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
28	2635	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
29	2659	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
30	2661	G/C	C/C	C/C	C/C	G/G	C/C	C/C	C/C
31	2704	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
32	2716	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
33	2846	G/G	G/G	A/G	G/G	G/G	G/A	G/A	G/G
34	3292	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
35	3470	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
36	4183	G/G	G/G	G/G	G/G	G/G	A/G	G/G	G/A
37	4201	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
38	4254	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
39	4384	A/C	C/C	C/C	C/C	A/A	C/C	C/C	C/C
40	4435	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
41	5180	G/C	C/C	C/C	C/C	G/C	C/C	C/C	C/C
42	5212	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C

(a) PS = polymorphic site;
(b) Position of PS in SEQ ID NO:1;
(c) Haplotype pairs are represented as 1st haplotype/2nd haplotype; with alleles of each haplotype shown 5′ to 3′ as 1st polymorphism/2nd polymorphism in each column;

and the frequency data in Tables 5 and 6.

34. A genome anthology for the Cytochrome P450, subfamily IID, Polypeptide 6 (CYP2D6) gene

which comprises two or more CYP2D6 isogenes selected from the group consisting of isogenes 1-34 shown in the table immediately below, and wherein each of the isogenes comprises the regions of SEQ ID NO:1 shown in the table immediately below and wherein each of the isogenes 1-34 is further defined by the corresponding sequence of polymorphisms whose positions and identities are set forth in the table immediately below:

Region	PS	PS	Isogene Number(d) (Part 1)									
Examined(a)	No.(b)	Position(c)	1	2	3	4	5	6	7	8	9	10
378-1363	1	636	A	G	G	G	G	G	G	G	G	G
378-1363	2	678	T	C	T	T	T	T	T	T	T	T
378-1363	3	769	G	G	C	G	G	G	G	G	G	G
378-1363	4	776	A	A	G	A	A	A	A	A	A	A
378-1363	5	825	G	G	G	A	G	G	G	G	G	G
378-1363	6	915	T	T	T	T	C	T	T	T	T	T
378-1363	7	1019	G	G	G	G	G	A	G	G	G	G
378-1363	8	1031	G	G	G	G	G	G	A	A	G	G
378-1363	9	1100	C	C	T	C	C	C	C	C	C	C
1701-2203	10	1827	G	G	G	G	G	G	G	G	C	G
1701-2203	11	1843	T	G	G	G	T	G	G	G	G	G
1701-2203	12	1966	G	G	G	G	G	G	G	G	G	A
1701-2203	13	1974	C	C	C	C	C	C	C	C	C	C
1701-2203	14	1984	A	A	A	A	A	A	A	A	A	A
1701-2203	15	1997	C	C	C	C	C	C	C	C	C	C
1701-2203	16	2014	T	T	T	T	T	T	T	T	T	T
1701-2203	17	2022	A	A	A	A	A	A	A	A	A	A
1701-2203	18	2023	C	C	C	C	C	C	C	C	T	C
1701-2203	19	2028	A	A	A	A	A	A	A	A	A	A
1701-2203	20	2036	T	T	T	T	T	T	T	T	T	T
1701-2203	21	2039	C	C	T	C	C	C	C	C	C	C
1701-2203	22	2062	A	A	A	A	A	A	A	A	A	A
1701-2203	23	2067	T	T	T	T	T	T	T	T	T	T
1701-2203	24	2118	C	C	C	C	C	C	C	C	C	C
1701-2203	25	2170	G	G	G	G	G	G	G	G	G	G
1701-2203	26	2179	G	G	G	G	G	G	G	G	G	G
2342-4555	27	2611	T	T	T	T	T	T	T	T	T	T
2342-4555	28	2635	T	T	T	T	T	T	T	T	C	T
2342-4555	29	2659	G	G	G	A	G	G	G	G	G	G
2342-4555	30	2661	G	C	C	C	G	C	C	G	C	C
2342-4555	31	2704	C	C	C	C	C	G	C	C	C	C
2342-4555	32	2716	G	G	G	G	G	G	G	G	G	G
2342-4555	33	2846	G	G	G	G	G	G	G	G	G	G
2342-4555	34	3292	G	G	G	A	G	G	G	G	G	G
2342-4555	35	3470	T	T	T	T	T	T	T	T	T	T
2342-4555	36	4183	G	G	G	A	G	G	G	G	G	G
2342-4555	37	4201	C	C	C	C	C	C	C	C	C	C
2342-4555	38	4254	T	T	T	T	T	T	T	T	T	T
2342-4555	39	4384	A	C	C	C	A	C	C	C	C	C
2342-4555	40	4435	C	C	C	C	C	C	C	C	C	C
4651-5435	41	5180	C	C	C	C	C	C	C	C	C	C
4651-5435	42	5212	C	C	C	C	C	C	C	C	C	C

Region Examined(a)	PS No.(b)	PS Position(c)	Isogene Number(d) (Part 2) 11	12	13	14	15	16	17	18	19	20
378-1363	1	636	G	G	G	G	G	G	G	G	G	G
378-1363	2	678	T	T	T	T	T	T	T	T	T	T
378-1363	3	769	G	G	G	G	G	G	G	G	G	G
378-1363	4	776	A	A	A	A	A	A	A	A	A	A
378-1363	5	825	G	G	G	G	G	G	G	G	G	G
378-1363	6	915	T	T	T	T	T	T	T	T	T	T
378-1363	7	1019	G	G	G	G	G	G	G	G	G	G
378-1363	8	1031	G	G	G	G	G	G	G	G	G	G
378-1363	9	1100	C	C	C	C	C	C	C	C	C	C
1701-2203	10	1827	G	G	G	G	G	G	G	G	G	G
1701-2203	11	1843	G	G	G	G	G	G	G	G	G	T
1701-2203	12	1966	G	G	G	G	G	G	G	G	G	G
1701-2203	13	1974	C	C	C	C	C	C	C	C	C	C
1701-2203	14	1984	A	A	A	A	A	A	A	A	A	A
1701-2203	15	1997	C	C	C	C	C	C	C	C	C	C
1701-2203	16	2014	T	T	T	T	T	T	T	T	T	T
1701-2203	17	2022	A	A	A	A	A	A	A	A	A	A
1701-2203	18	2023	C	C	C	C	C	C	C	T	T	C
1701-2203	19	2028	A	A	A	A	A	A	A	A	A	A
1701-2203	20	2036	T	T	T	T	T	T	T	T	T	T
1701-2203	21	2039	C	C	C	C	C	C	C	C	C	C
1701-2203	22	2062	A	A	A	A	A	A	A	A	A	A
1701-2203	23	2067	G	G	T	T	T	T	T	T	T	T
1701-2203	24	2118	C	C	C	C	C	C	C	C	C	C
1701-2203	25	2170	G	G	G	G	G	G	G	G	G	A
1701-2203	26	2179	C	G	G	G	G	G	G	G	G	G
2342-4555	27	2611	T	T	A	T	T	T	T	T	T	T
2342-4555	28	2635	T	T	T	T	T	T	T	T	T	T
2342-4555	29	2659	G	G	G	G	G	G	G	G	G	G
2342-4555	30	2661	G	G	C	C	C	G	G	C	C	G
2342-4555	31	2704	C	C	C	C	C	C	C	C	C	C
2342-4555	32	2716	G	G	G	A	G	G	G	G	G	G
2342-4555	33	2846	G	G	G	G	G	G	G	G	G	G
2342-4555	34	3292	G	G	G	G	G	G	G	G	G	G
2342-4555	35	3470	T	T	T	T	T	T	T	C	T	T
2342-4555	36	4183	G	G	G	G	G	G	G	G	G	G
2342-4555	37	4201	C	C	C	C	C	C	C	C	C	C
2342-4555	38	4254	T	T	T	T	T	T	T	T	T	T
2342-4555	39	4384	A	A	C	C	C	A	A	C	C	A
2342-4555	40	4435	C	C	C	C	C	C	C	C	C	C
4651-5435	41	5180	G	G	C	C	C	C	G	C	C	G
4651-5435	42	5212	C	C	C	C	C	C	C	C	C	C

Region Examined(a)	PS No.(b)	PS Position(c)	Isogene Number(d) (Part 3) 21	22	23	24	25	26	27	28	29	30
378-1363	1	636	G	G	G	G	G	G	G	G	G	G
378-1363	2	678	T	T	T	T	T	T	T	T	T	T
378-1363	3	769	G	G	G	G	G	G	G	G	G	G
378-1363	4	776	A	A	A	A	A	A	A	A	A	A
378-1363	5	825	G	G	G	G	G	G	G	G	G	G
378-1363	6	915	T	T	T	T	T	T	T	T	T	T
378-1363	7	1019	G	G	G	G	G	G	G	G	G	G
378-1363	8	1031	G	G	G	G	G	G	G	G	G	G
378-1363	9	1100	C	C	C	C	T	T	T	T	T	T
1701-2203	10	1827	G	G	G	G	G	G	G	G	G	G
1701-2203	11	1843	T	T	T	T	G	G	G	G	G	G
1701-2203	12	1966	G	G	G	G	G	G	G	G	G	G
1701-2203	13	1974	C	C	C	C	A	A	A	A	A	A
1701-2203	14	1984	A	A	A	A	A	A	A	A	G	G
1701-2203	15	1997	C	C	C	C	C	G	G	G	G	G
1701-2203	16	2014	T	T	T	T	T	T	T	T	T	T
1701-2203	17	2022	A	A	A	A	A	A	A	A	A	A
1701-2203	18	2023	C	C	C	C	C	C	C	C	C	C
1701-2203	19	2028	A	A	A	A	A	A	A	A	A	A
1701-2203	20	2036	T	T	T	T	T	T	T	T	T	T
1701-2203	21	2039	C	C	C	C	C	C	C	C	C	C
1701-2203	22	2062	A	A	A	A	A	A	A	A	A	A
1701-2203	23	2067	T	T	T	T	T	T	T	T	T	T
1701-2203	24	2118	C	C	C	T	C	C	C	C	C	C
1701-2203	25	2170	G	G	G	G	G	G	G	G	G	G
1701-2203	26	2179	G	G	G	G	G	G	G	G	G	G
2342-4555	27	2611	T	T	T	T	T	T	T	T	T	T
2342-4555	28	2635	T	T	T	T	T	T	T	T	T	T
2342-4555	29	2659	G	G	G	G	G	G	G	G	G	G
2342-4555	30	2661	C	G	G	G	C	C	C	C	C	C
2342-4555	31	2704	C	C	C	C	C	C	C	C	C	C
2342-4555	32	2716	G	G	G	G	G	G	G	G	G	G
2342-4555	33	2846	G	G	G	G	G	A	G	G	A	G
2342-4555	34	3292	G	G	G	G	G	G	G	G	G	G
2342-4555	35	3470	T	T	T	T	T	T	T	T	T	T
2342-4555	36	4183	G	G	G	G	G	G	G	G	G	G
2342-4555	37	4201	C	C	C	C	C	C	C	C	C	C
2342-4555	38	4254	T	T	T	T	T	T	T	T	T	T
2342-4555	39	4384	A	A	A	A	C	C	C	C	C	C
2342-4555	40	4435	C	C	C	C	C	C	A	C	C	A
4651-5435	41	5180	G	C	G	C	C	C	C	C	C	C
4651-5435	42	5212	C	C	C	C	C	C	C	C	C	C

Region Examined(a)	PS No.(b)	PS Position(c)	Isogene Number(d) (Part 4) 31	32	33	34
378-1363	1	636	G	G	G	G
378-1363	2	678	T	T	T	T
378-1363	3	769	G	G	G	G
378-1363	4	776	A	A	A	A
378-1363	5	825	G	G	G	G
378-1363	6	915	T	T	T	T
378-1363	7	1019	G	G	G	G
378-1363	8	1031	G	G	G	G
378-1363	9	1100	T	T	T	T
1701-2203	10	1827	G	G	G	G
1701-2203	11	1843	G	G	G	T
1701-2203	12	1966	G	G	G	G
1701-2203	13	1974	C	C	C	A
1701-2203	14	1984	A	A	A	A
1701-2203	15	1997	C	C	C	G
1701-2203	16	2014	T	T	T	T
1701-2203	17	2022	A	A	A	A
1701-2203	18	2023	C	C	C	C
1701-2203	19	2028	A	A	A	A
1701-2203	20	2036	T	T	T	T
1701-2203	21	2039	T	T	T	C
1701-2203	22	2062	A	A	A	A
1701-2203	23	2067	T	T	T	T
1701-2203	24	2118	C	C	C	C
1701-2203	25	2170	G	G	G	G
1701-2203	26	2179	G	G	G	G
2342-4555	27	2611	T	T	T	T
2342-4555	28	2635	T	T	T	T
2342-4555	29	2659	G	G	G	G
2342-4555	30	2661	C	C	C	C
2342-4555	31	2704	C	C	C	C
2342-4555	32	2716	G	G	G	G
2342-4555	33	2846	A	G	G	A
2342-4555	34	3292	G	G	G	G
2342-4555	35	3470	T	T	T	T
2342-4555	36	4183	G	G	G	G
2342-4555	37	4201	C	C	T	C
2342-4555	38	4254	T	T	T	T
2342-4555	39	4384	C	C	C	A
2342-4555	40	4435	C	C	C	C
4651-5435	41	5180	C	C	C	G
4651-5435	42	5212	C	C	C	C

(a) Region examined represents the nucleotide positions defining the start and stop positions within SEQ ID NO:1 of the regions sequenced;
(b) PS = polymorphic site;
(c) Position of PS within SEQ ID NO:1;
(d) Alleles for CYP2D6 isogenes are presented 5′ to 3′ in each column.

POLYMORPHISMS IN THE CYP2D6 GENE

```
GACCCTCAGG CAGCCCGGGA GGATGTTGTC ACAGGCTGGG GCAAGGGCCT
TCCGGCTACC AACTGGGAGC TCTGGGAACA GCCCTGTTGC AAACAAGAAG      100
CCATAGCCCG GCCAGAGCCC AGGAATGTGG GCTGGGCTGG GAGCAGCCTC
TGGACAGGAG TGGTCCCATC CAGGAAACCT CCGGCATGGC TGGGAAGTGG      200
GGTACTTGGT GCCGGGTCTG TATGTGTGTG TGACTGGTGT GTGTGAGAGA
GAATGTGTGC CCTAAGTGTC AGTGTGAGTC TGTGTATGTG TGAATATTGT      300
CTTTGTGTGG GTGATTTTCT GCGTGTGTAA TCGTGTCCCT GCAAGTGTGA
ACAAGTGGAC AAGTGTCTGG GAGTGGACAA GAGATCTGTG CACCATCAGG      400
TGTGTGCATA GCGTCTGTGC ATGTCAAGAG TGCAAGGTGA AGTGAAGGGA
CCAGGCCCAT GATGCCACTC ATCATCAGGA GCTCTAAGGC CCCAGGTAAG      500
TGCCAGTGAC AGATAAGGGT GCTGAAGGTC ACTCTGGAGT GGGCAGGTGG
GGGTAGGGAA AGGGCAAGGC CATGTTCTGG AGGAGGGGTT GTGACTACAT      600
TAGGGTGTAT GAGCCTAGCT GGGAGGTGGA TGGCCGGGTC CACTGAAACC
                                  A
CTGGTTATCC CAGAAGGCTT TGCAGGCTTC AGGAGCTTGG AGTGGGGAGA      700
                               C
GGGGGTGACT TCTCCGACCA GGCCCCTCCA CCGGCCTACC CTGGGTAAGG
GCCTGGAGCA GGAAGCAGGG GCAAGAACCT CTGGAGCAGC CCATACCCGC      800
                    C       G
CCTGGCCTGA CTCTGCCACT GGCAGCACAG TCAACACAGC AGGTTCACTC
                              A
ACAGCAGAGG GCAAAGGCCA TCATCAGCTC CCTTTATAAG GGAAGGGTCA      900
CGCGCTCGGT GTGCTGAGAG TGTCCTGCCT GGTCCTCTGT GCCTGGTGGG
                  C
GTGGGGGTGC CAGGTGTGTC CAGAGGAGCC CATTTGGTAG TGAGGCAGGT     1000
ATGGGGCTAG AAGCACTGGT GCCCCTGGCC GTGATAGTGG CCATCTTCCT
                    A                 A
      [exon 1: 1001..
GCTCCTGGTG GACCTGATGC ACCGGCGCCA ACGCTGGGCT GCACGCTACC     1100
                                                      T
CACCAGGCCC CCTGCCACTG CCCGGGCTGG GCAACCTGCT GCATGTGGAC
TTCCAGAACA CACCATACTG CTTCGACCAG GTGAGGGAGG AGGTCCTGGA     1200
                   ..1180]
GGGCGGCAGA GGTGCTGAGG CTCCCCTACC AGAAGCAAAC ATGGATGGTG
GGTGAAACCA CAGGCTGGAC CAGAAGCCAG GCTGAGAAGG GGAAGCAGGT     1300
TTGGGGGACG TCCTGGAGAA GGGCATTTAT ACATGGCATG AAGGACTGGA
TTTTCCAAAG GCCAAGGAAG AGTAGGGCAA GGGCCTGGAG GTGGAGCTGG     1400
ACTTGGCAGT GGGCATGCAA GCCCATTGGG CAACATATGT TATGGAGTAC
AAAGTCCCTT CTGCTGACAC CAGAAGGAAA GGCCTTGGGA ATGGAAGATG     1500
AGTTAGTCCT GAGTGCCGTT TAAATCACGA AATCGAGGAT GAAGGGGGTG
CAGTGACCCG GTTCAAACCT TTTGCACTGT GGGTCCTCGG GCCTCACTGC     1600
CTCACCGGCA TGGACCATCA TCTGGGAATG GGATGCTAAC TGGGGCCTCT
CGGCAATTTT GGTGACTCTT GCAAGGTCAT ACCTGGGTGA CGCATCCAAA     1700
CTGAGTTCCT CCATCACAGA AGGTGTGACC CCCACCCCCG CCCCACGATC
AGGAGGCTGG GTCTCCTCCT TCCACCTGCT CACTCCTGGT AGCCCCGGGG     1800
GTCGTCCAAG GTTCAAATAG GACTAGGACC TGTAGTCTGG GGTGATCCTG
                              C                   G
GCTTGACAAG AGGCCCTGAC CCTCCCTCTG CAGTTGCGGC GCCGCTTCGG     1900
      [exon 2: 1884..
GGACGTGTTC AGCCTGCAGC TGGCCTGGAC GCCGGTGGTC GTGCTCAATG
GGCTGGCGGC CGTGCGCGAG GCGCTGGTGA CCCACGGCGA GGACACCGCC     2000
                A         A         G                 G
GACCGCCCGC CTGTGCCCAT CACCCAGATC CTGGGTTTCG GGCCGCGTTC
                C         TT    G         C  T
```

FIGURE 1A

```
CCAAGGCAAG CAGCGGTGGG GACAGAGACA GATTTCCGTG GGACCCGGGT      2100
            G    G
                  ..2055]
GGGTGATGAC CGTAGTCCGA GCTGGGCAGA GAGGGCGCGG GGTCGTGGAC
                 T
ATGAAACAGG CCAGCGAGTG GGGACAGCGG GCCAAGAAAC CACCTGCACT      2200
                   A          C
AGGGAGGTGT GAGCATGGGG ACGAGGGCGG GGCTTGTGAC GAGTGGGCGG
GGCCACTGCC GAGACCTGGC AGGAGCCCAA TGGGTGAGCG TGGCGCATTT      2300
CCCAGCTGGA ATCCGGTGTC GAAGTGGGGG CGGGGACCGC ACCTGTGCTG
TAAGCTCAGT GTGGGTGGCG CGGGGCCCGC GGGGTCTTCC CTGAGTGCAA      2400
AGGCGGTCAG GGTGGGCAGA GACGAGGTGG GGCAAAGCCT GCCCCAGCCA
AGGGAGCAAG GTGGATGCAC AAAGAGTGGG CCCTGTGACC AGCTGGACAG      2500
AGCCAGGGAC TGCGGGAGAC CAGGGGGAGC ATAGGGTTGG AGTGGGTGGT
GGATGGTGGG GCTAATGCCT TCATGGCCAC GCGCACGTGC CCGTCCCACC      2600
CCCAGGGGTG TTCCTGGCGC GCTATGGGCC CGCGTGGCGC GAGCAGAGGC
           A                          C
      [exon 3: 2606..
GCTTCTCCGT GTCCACCTTG CGCAACTTGG GCCTGGGCAA GAAGTCGCTG      2700
         A  C
GAGCAGTGGG TGACCGAGGA GGCCGCCTGC CTTTGTGCCG CCTTCGCCAA
   G             A
CCACTCCGGT GGGTGATGGG CAGAAGGGCA CAAAGCGGGA ACTGGGAAGG      2800
                 ..2758]
CGGGGGACGG GGAAGGCGAC CCCTTACCCG CATCTCCCAC CCCCAGGACG
                                                    A
      [exon 4: 2847..
CCCCTTTCGC CCCAACGGTC TCTTGGACAA AGCCGTGAGC AACGTGATCG      2900
CCTCCCTCAC CTGCGGGCGC CGCTTCGAGT ACGACGACCC TCGCTTCCTC
AGGCTGCTGG ACCTAGCTCA GGAGGGACTG AAGGAGGAGT CGGGCTTTCT      3000
GCGCGAGGTG CGGAGCGAGA GACCGAGGAG TCTCTGCAGG GCGAGCTCCC
                 ..3007]
GAGAGGTGCC GGGGCTGGAC TGGGGCCTCG GAAGAGCAGG ATTTGCATAG      3100
ATGGGTTTGG GAAAGGACAT TCCAGGAGAC CCCACTGTAA GAAGGGCCTG
GAGGAGGAGG GGACATCTCA GACATGGTCG TGGGAGAGGT GTGCCCGGGT      3200
CAGGGGGCAC CAGGAGAGGC CAAGGACTCT GTACCTCCTA TCCACGTCAG
AGATTTCGAT TTTAGGTTTC TCCTCTGGGC AAGGAGAGAG GGTGGAGGCT      3300
                                              A
GGCACTTGGG GAGGGACTTG GTGAGGTCAG TGGTAAGGAC AGGCAGGCCC
TGGGTCTACC TGGAGATGGC TGGGGCCTGA GACTTGTCCA GGTGAACGCA      3400
GAGCACAGGA GGGATTGAGA CCCCGTTCTG TCTGGTGTAG GTGCTGAATG
      [exon 5: 3441..
CTGTCCCCGT CCTCCTGCAT ATCCCAGCGC TGGCTGGCAA GGTCCTACGC      3500
                    C
TTCCAAAAGG CTTTCCTGAC CCAGCTGGAT GAGCTGCTAA CTGAGCACAG
GATGACCTGG GACCCAGCCC AGCCCCCCCG AGACCTGACT GAGGCCTTCC      3600
TGGCAGAGAT GGAGAAGGTG AGAGTGGCTG CCACGGTGGG GGGCAAGGGT
                   ..3617]
GGTGGGTTGA GCGTCCCAGG AGGAATGAGG GGAGGCTGGG CAAAAGGTTG      3700
GACCAGTGCA TCACCCGGCG AGCCGCATCT GGGCTGACAG GTGCAGAATT
GGAGGTCATT TGGGGGCTAC CCCGTTCTGT CCCGAGTATG CTCTCGGCCC      3800
TGCTCAGGCC AAGGGGAACC CTGAGAGCAG CTTCAATGAT GAGAACCTGC
       [exon 6: 3808..
GCATAGTGGT GGCTGACCTG TTCTCTGCCG GGATGGTGAC CACCTCGACC      3900
ACGCTGGCCT GGGGCCTCCT GCTCATGATC CTACATCCGG ATGTGCAGCG
                   ..3949]
TGAGCCCATC TGGGAAACAG TGCAGGGGCC GAGGGAGGAA GGGTACAGGC      4000
GGGGGCCCAT GAACTTTGCT GGGACACCCG GGGCTCCAAG CACAGGCTTG
ACCAGGATCC TGTAAGCCTG ACCTCCTCCA ACATAGGAGG CAAGAAGGAG      4100
```

FIGURE 1B

```
TGTCAGGGCC GGACCCCCTG GGTGCTGACC CATTGTGGGG ACGCATGTCT
GTCCAGGCCG TGTCCAACAG GAGATCGACG ACGTGATAGG GCAGGTGCGG      4200
                                  A
      [exon 7: 4157..
CGACCAGAGA TGGGTGACCA GGCTCACATG CCCTACACCA CTGCCGTGAT
T
TCATGAGGTG CAGCGCTTTG GGGACATCGT CCCCCTGGGT GTGACCCATA      4300
   C
TGACATCCCG TGACATCGAA GTACAGGGCT TCCGCATCCC TAAGGTAGGC
                  ..4344]
CTGGCGCCCT CCTCACCCCA GCTCAGCACC AGCACCTGGT GATAGCCCCA      4400
                                  C
GCATGGCTAC TGCCAGGTGG GCCCACTCTA GGAACCCTGG CCACCTAGTC
                                  A
CTCAATGCCA CCACACTGAC TGTCCCCACT TGGGTGGGGG GTCCAGAGTA      4500
TAGGCAGGGC TGGCCTGTCC ATCCAGAGCC CCCGTCTAGT GGGGAGACAA
ACCAGGACCT GCCAGAATGT TGGAGGACCC AACGCCTGCA GGGAGAGGGG      4600
GCAGTGTGGG TGCCTCTGAG AGGTGTGACT GCGCCCTGCT GTGGGGTCGG
AGAGGGTACT GTGGAGCTTC TCGGGCGCAG GACTAGTTGA CAGAGTCCAG      4700
CTGTGTGCCA GGCAGTGTGT GTCCCCCGTG TGTTTGGTGG CAGGGGTCCC
AGCATCCTAG AGTCCAGTCC CCACTCTCAC CCTGCATCTC CTGCCCAGGG      4800
      [exon 8: 4799..
AACGACACTC ATCACCAACC TGTCATCGGT GCTGAAGGAT GAGGCCGTCT
GGGAGAAGCC CTTCCGCTTC CACCCCGAAC ACTTCCTGGA TGCCCAGGGC      4900
CACTTTGTGA AGCCGGAGGC CTTCCTGCCT TTCTCAGCAG GTGCCTGTGG
                  ..4940]
GGAGCCCGGC TCCCTGTCCC CTTCCGTGGA GTCTTGCAGG GGTATCACCC      5000
AGGAGCCAGG CTCACTGACG CCCCTCCCCT CCCCACAGGC CGCCGTGCAT
      [exon 9: 5039..
GCCTCGGGGA GCCCCTGGCC CGCATGGAGC TCTTCCTCTT CTTCACCTCC      5100
CTGCTGCAGC ACTTCAGCTT CTCGGTGCCC ACTGGACAGC CCCGGCCCAG
CCACCATGGT GTCTTTGCTT TCCTGGTGAG CCCATCCCCC TATGAGCTTT      5200
                                  C
GTGCTGTGCC CCGCTAGAAT GGGGTACCTA GTCCCCAGCC TGCTCCCTAG
          T
                  ..5217]
CCAGAGGCTC TAATGTACAA TAAAGCAATG TGGTAGTTCC AACTCGGGTC      5300
CCCTGCTCAC GCCCTCGTTG GGATCATCCT CCTCAGGGCA ACCCCACCCC
TGCCTCATTC CTGCTTACCC CACCGCCTGG CCGCATTTGA GACAGGGGTA      5400
CGTTGAGGCT GAGCAGATGT CAGTTACCCT TGCCCATAAT CCCATGTCCC
CCACTGACCC AACTCTGACT GCCCAGATTG GTGACAAGGA CTACATTGTC      5500
CTGGCATGTG GGGAAGGGGC CAGAATGGGC TGACTAGAGG TGTCAGTCAG
CCCTGGATGT GGTGGAGAGG GCAGGACTCA GCCTGGAGGC CCATATTTCA      5600
GGCCTAACTC AGCCCACCCC ACATCAGGGA CAGCAGTCCT GCCAGCACCA
TCACAACAGT CACCTCCCTT CATATATGAC ACCCCAAAAC GGAAGACAAA      5700
TCATGGCGTC AGGGAGCTAT ATGCCAGGGC TACCTACCTC CCAGGGCTCA
GTCGGCAGGT GCCAGAACGT TCCCTGGGAA GGCCCCATGG AAGCCCAGGA      5800
CTGAGCCACC ACCCTCAGCC TCGTCACCTC ACCACAGGAC TGGCTACCTC
TCTGGGCCCT CAGGGATGCT GCTGTACAGA CCCCTGACCA GTGACGAGTT      5900
CGCACTCAGG GCCAGGCTGG CGCTGGAGGA GGACACTTGT TTGGCTCCAA
CCCTAGGTAC CATCCTCCCA GTAGGGATCA GGCAGGGCCC ACAGGCCTGC      6000
CCTAGGGACA GGAGTCAACC TTGGACCCAT AAGGCACTGG GGCGGGCAGA
GAAGGAGGAG GTGGCATGGG CAGCTGAGAG CCAGAGACCC TGACCCTAGT      6100
CCTTGCTCTG CCATTACCCC GTGTGACCCC GGGCCCACCC TTCCCCACCC
TTCCCCACCC CGGGCTTCTG TTTCCTTCTG CCAACGAGAA GGCTGCTTCA      6200
CCTGCCCCGA GTCCTGTCTT CCTGCTCTGC CTTCTGGGGC TGTGGCCCTT
GCTGGCCTGG AGCCCCAACC AAGGGCAGGG ACTGCTGTCC TCCACGTCTG      6300
TCCTCACCGA CATAATGGGC TGGGCTGGGC ACACAGGCAG TGCCCAAGAG
TTTCTAATGA GCATATGATT ACCTGAGTCC TGGGCAGACC TTCTTAGGGA      6400
```

FIGURE 1C

ACAGCCTGGG ACAGAGAACC ACAGACACTC TGAGGAGCCA CCCTGAGGCC
TCTTTTGCCA GAGGACCCTA CA 6472

FIGURE 1D

POLYMORPHISMS IN THE CODING SEQUENCE OF CYP2D6

```
ATGGGGCTAG AAGCACTGGT GCCCCTGGCC GTGATAGTGG CCATCTTCCT
                  A               A
GCTCCTGGTG GACCTGATGC ACCGGCGCCA ACGCTGGGCT GCACGCTACC      100
                                                    T
CACCAGGCCC CCTGCCACTG CCCGGGCTGG GCAACCTGCT GCATGTGGAC
TTCCAGAACA CACCATACTG CTTCGACCAG TTGCGGCGCC GCTTCGGGGA      200
CGTGTTCAGC CTGCAGCTGG CCTGGACGCC GGTGGTCGTG CTCAATGGGC
TGGCGGCCGT GCGCGAGGCG CTGGTGACCC ACGGCGAGGA CACCGCCGAC      300
             A         A           G              G
CGCCCGCCTG TGCCCATCAC CCAGATCCTG GGTTTCGGGC CGCGTTCCCA
           C       TT     G         C  T
AGGGGTGTTC CTGGCGCGCT ATGGGCCCGC GTGGCGCGAG CAGAGGCGCT      400
       A                             C
TCTCCGTGTC CACCTTGCGC AACTTGGGCC TGGGCAAGAA GTCGCTGGAG
     A C
CAGTGGGTGA CCGAGGAGGC CGCCTGCCTT TGTGCCGCCT TCGCCAACCA      500
G            A
CTCCGGACGC CCCTTTCGCC CCAACGGTCT CTTGGACAAA GCCGTGAGCA
ACGTGATCGC CTCCCTCACC TGCGGGCGCC GCTTCGAGTA CGACGACCCT      600
CGCTTCCTCA GGCTGCTGGA CCTAGCTCAG GAGGGACTGA AGGAGGAGTC
GGGCTTTCTG CGCGAGGTGC TGAATGCTGT CCCCGTCCTC CTGCATATCC      700
                                                  C
CAGCGCTGGC TGGCAAGGTC CTACGCTTCC AAAAGGCTTT CCTGACCCAG
CTGGATGAGC TGCTAACTGA GCACAGGATG ACCTGGGACC CAGCCCAGCC      800
CCCCCGAGAC CTGACTGAGG CCTTCCTGGC AGAGATGGAG AAGGCCAAGG
GGAACCCTGA GAGCAGCTTC AATGATGAGA ACCTGCGCAT AGTGGTGGCT      900
GACCTGTTCT CTGCCGGGAT GGTGACCACC TCGACCACGC TGGCCTGGGG
CCTCCTGCTC ATGATCCTAC ATCCGGATGT GCAGCGCCGT GTCCAACAGG     1000
AGATCGACGA CGTGATAGGG CAGGTGCGGC GACCAGAGAT GGGTGACCAG
            A                   T
GCTCACATGC CCTACACCAC TGCCGTGATT CATGAGGTGC AGCGCTTTGG     1100
                                  C
GGACATCGTC CCCCTGGGTG TGACCCATAT GACATCCCGT GACATCGAAG
TACAGGGCTT CCGCATCCCT AAGGGAACGA CACTCATCAC CAACCTGTCA     1200
TCGGTGCTGA AGGATGAGGC CGTCTGGGAG AAGCCCTTCC GCTTCCACCC
CGAACACTTC CTGGATGCCC AGGGCCACTT TGTGAAGCCG GAGGCCTTCC     1300
TGCCTTTCTC AGCAGGCCGC CGTGCATGCC TCGGGGAGCC CCTGGCCCGC
ATGGAGCTCT TCCTCTTCTT CACCTCCCTG CTGCAGCACT TCAGCTTCTC     1400
GGTGCCCACT GGACAGCCCC GGCCCAGCCA CCATGGTGTC TTTGCTTTCC
TGGTGAGCCC ATCCCCCTAT GAGCTTTGTG CTGTGCCCCG CTAG           1494
      C                                   T
```

FIGURE 2

ISOFORMS OF THE CYP2D6 PROTEIN

```
MGLEALVPLA VIVAIFLLLV DLMHRRQRWA ARYPPGPLPL PGLGNLLHVD
      M    M                          S
FQNTPYCFDQ LRRRFGDVFS LQLAWTPVVV LNGLAAVREA LVTHGEDTAD      100
                                        H   M  R
RPPVPITQIL GFGPRSQGVF LARYGPAWRE QRRFSVSTLR NLGLGKKSLE
   A  F V            I         R        I
QWVTEEAACL CAAFANHSGR PFRPNGLLDK AVSNVIASLT CGRRFEYDDP      200
E   K
RFLRLLDLAQ EGLKEESGFL REVLNAVPVL LHIPALAGKV LRFQKAFLTQ
LDELLTEHRM TWDPAQPPRD LTEAFLAEME KAKGNPESSF NDENLRIVVA      300
DLFSAGMVTT STTLAWGLLL MILHPDVQRR VQQEIDDVIG QVRRPEMGDQ
                                      M      !
AHMPYTTAVI HEVQRFGDIV PLGVTHMTSR DIEVQGFRIP KGTTLITNLS      400
SVLKDEAVWE KPFRFHPEHF LDAQGHFVKP EAFLPFSAGR RACLGEPLAR
MELFLFFTSL LQHFSFSVPT GQPRPSHHGV FAFLVSPSPY ELCAVPR         497
                                    T           C
```

FIGURE 3

SEQ Listing CYP2D6.ST25
SEQUENCE LISTING

<110> Genaissance Pharmaceuticals, Inc.
Anastasio, Alison E
Chew, Anne
Choi, Julie Y
Denton, R. Rex
Nandabalan, Krishnan
Peterson, Nathan
Rounds, Eileen

<120> HAPLOTYPES OF THE CYP2D6 GENE

<130> CYP2D6-MWH_0398-PCT

<140> TBA
<141> 2001-11-09

<150> 60/247,943
<151> 2000-11-09

<160> 151

<170> PatentIn version 3.1

<210> 1
<211> 6472
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (636)..(636)
<223> PS1: polymorphic base G or A

<220>
<221> allele
<222> (678)..(678)
<223> PS2 : polymorphic base T or C

<220>
<221> allele
<222> (769)..(769)
<223> PS3 : polymorphic base G or C

<220>
<221> allele
<222> (776)..(776)
<223> PS4: polymorphic base A or G

<220>
<221> allele
<222> (825)..(825)
<223> PS5: polymorphic base G or A

<220>
<221> allele
<222> (915)..(915)
<223> PS6: polymorphic base T or C

<220>
<221> allele

<222> (1019)..(1019)
<223> PS7: polymorphic base G or A

<220>
<221> allele
<222> (1031)..(1031)
<223> PS8: polymorphic base G or A

<220>
<221> allele
<222> (1100)..(1100)
<223> PS9: polymorphic base C or T

<220>
<221> allele
<222> (1827)..(1827)
<223> PS10: polymorphic base G or C

<220>
<221> allele
<222> (1843)..(1843)
<223> PS11: polymorphic base T or G

<220>
<221> allele
<222> (1966)..(1966)
<223> PS12: polymorphic base G or A

<220>
<221> allele
<222> (1974)..(1974)
<223> PS13: polymorphic base C or A

<220>
<221> allele
<222> (1984)..(1984)
<223> PS14: polymorphic base A or G

<220>
<221> allele
<222> (1997)..(1997)
<223> PS15: polymorphic base C or G

<220>
<221> allele
<222> (2014)..(2014)
<223> PS16: polymorphic base T or C

<220>
<221> allele
<222> (2022)..(2022)
<223> PS17: polymorphic base A or T

<220>
<221> allele
<222> (2023)..(2023)
<223> PS18: polymorphic base C or T

SEQ Listing CYP2D6.ST25

<220>
<221> allele
<222> (2028)..(2028)
<223> PS19: polymorphic base A or G

<220>
<221> allele
<222> (2036)..(2036)
<223> PS20: polymorphic base T or C

<220>
<221> allele
<222> (2039)..(2039)
<223> PS21: polymorphic base C or T

<220>
<221> allele
<222> (2062)..(2062)
<223> PS22: polymorphic base A or G

<220>
<221> allele
<222> (2067)..(2067)
<223> PS23: polymorphic base Tor G

<220>
<221> allele
<222> (2118)..(2118)
<223> PS24: polymorphic base C or T

<220>
<221> allele
<222> (2170)..(2170)
<223> PS25: polymorphic base G or A

<220>
<221> allele
<222> (2179)..(2179)
<223> PS26: polymorphic base G or C

<220>
<221> allele
<222> (2611)..(2611)
<223> PS27: polymorphic base T or A

<220>
<221> allele
<222> (2635)..(2635)
<223> PS28: polymorphic base T or C

<220>
<221> allele
<222> (2659)..(2659)
<223> PS29: polymorphic base G or A

SEQ Listing CYP2D6.ST25

<220>
<221> allele
<222> (2661)..(2661)
<223> PS30: polymorphic base G or C

<220>
<221> allele
<222> (2704)..(2704)
<223> PS31: polymorphic base C or G

<220>
<221> allele
<222> (2716)..(2716)
<223> PS32: polymorphic base G or A

<220>
<221> allele
<222> (2846)..(2846)
<223> PS33: polymorphic base G or A

<220>
<221> allele
<222> (3292)..(3292)
<223> PS34: polymorphic base G or A

<220>
<221> allele
<222> (3470)..(3470)
<223> PS35: polymorphic base T or C

<220>
<221> allele
<222> (4183)..(4183)
<223> PS36: polymorphic base G or A

<220>
<221> allele
<222> (4201)..(4201)
<223> PS37: polymorphic base C or T

<220>
<221> allele
<222> (4254)..(4254)
<223> PS38: polymorphic base T or C

<220>
<221> allele
<222> (4384)..(4384)
<223> PS39: polymorphic base A or C

<220>
<221> allele
<222> (4435)..(4435)
<223> PS40: polymorphic base C or A

<220>
<221> allele

<222> (5180)..(5180)
<223> PS41: polymorphic base G or C

<220>
<221> allele
<222> (5212)..(5212)
<223> PS42: polymorphic base C or T

<220>
<221> allele
<222> (5212)..(5212)
<223> PS42: polymorphic base C or T

<400> 1

gaccctcagg	cagcccggga	ggatgttgtc	acaggctggg	gcaagggcct	tccggctacc	60
aactgggagc	tctgggaaca	gccctgttgc	aaacaagaag	ccatagcccg	gccagagccc	120
aggaatgtgg	gctgggctgg	gagcagcctc	tggacaggag	tggtcccatc	caggaaacct	180
ccggcatggc	tgggaagtgg	ggtacttggt	gccgggtctg	tatgtgtgtg	tgactggtgt	240
gtgtgagaga	gaatgtgtgc	cctaagtgtc	agtgtgagtc	tgtgtatgtg	tgaatattgt	300
ctttgtgtgg	gtgattttct	gcgtgtgtaa	tcgtgtccct	gcaagtgtga	acaagtggac	360
aagtgtctgg	gagtggacaa	gagatctgtg	caccatcagg	tgtgtgcata	gcgtctgtgc	420
atgtcaagag	tgcaaggtga	agtgaaggga	ccaggcccat	gatgccactc	atcatcagga	480
gctctaaggc	cccaggtaag	tgccagtgac	agataagggt	gctgaaggtc	actctggagt	540
gggcaggtgg	gggtagggaa	agggcaaggc	catgttctgg	aggaggggtt	gtgactacat	600
tagggtgtat	gagcctagct	gggaggtgga	tggccrggtc	cactgaaacc	ctggttatcc	660
cagaaggctt	tgcaggcytc	aggagcttgg	agtggggaga	gggggtgact	tctccgacca	720
ggcccctcca	ccggcctacc	ctgggtaagg	gcctggagca	ggaagcagsg	gcaagracct	780
ctggagcagc	ccatacccgc	cctggcctga	ctctgccact	ggcarcacag	tcaacacagc	840
aggttcactc	acagcagagg	gcaaaggcca	tcatcagctc	cctttataag	ggaagggtca	900
cgcgctcggt	gtgcygagag	tgtcctgcct	ggtcctctgt	gcctggtggg	gtgggggtgc	960
caggtgtgtc	cagaggagcc	catttggtag	tgaggcaggt	atggggctag	aagcactgrt	1020
gcccctggcc	rtgatagtgg	ccatcttcct	gctcctggtg	gacctgatgc	accggcgcca	1080
acgctgggct	gcacgctacy	caccaggccc	cctgccactg	cccgggctgg	gcaacctgct	1140
gcatgtggac	ttccagaaca	caccatactg	cttcgaccag	gtgagggagg	aggtcctgga	1200
gggcggcaga	ggtgctgagg	ctcccctacc	agaagcaaac	atggatggtg	ggtgaaacca	1260
caggctggac	cagaagccag	gctgagaagg	ggaagcaggt	ttgggggacg	tcctggagaa	1320
gggcatttat	acatggcatg	aaggactgga	ttttccaaag	gccaaggaag	agtagggcaa	1380
gggcctggag	gtggagctgg	acttggcagt	gggcâtgcaa	gcccattggg	caacatatgt	1440
tatggagtac	aaagtccctt	ctgctgacac	cagaaggaaa	ggccttggga	atggaagatg	1500
agttagtcct	gagtgccgtt	taaatcacga	aatcgaggat	gaaggggtg	cagtgacccg	1560
gttcaaacct	tttgcactgt	gggtcctcgg	gcctcactgc	ctcaccggca	tggaccatca	1620
tctgggaatg	ggatgctaac	tggggcctct	cggcaatttt	ggtgactctt	gcaaggtcat	1680
acctgggtga	cgcatccaaa	ctgagttcct	ccatcacaga	aggtgtgacc	cccaccccg	1740
ccccacgatc	aggaggctgg	gtctcctcct	tccacctgct	cactcctggt	agccccgggg	1800
gtcgtccaag	gttcaaatag	gactagsacc	tgtagtctgg	ggkgatcctg	gcttgacaag	1860
aggccctgac	cctccctctg	cagttgcggc	gccgcttcgg	ggacgtgttc	agcctgcagc	1920
tggcctggac	gccggtggtc	gtgctcaatg	ggctggcggc	cgtgcrcgag	gcgmtggtga	1980
cccrcggcga	ggacacsgcc	gaccgcccgc	ctgygcccat	cwyccagrtc	ctgggyttyg	2040
ggccgcgttc	ccaaggcaag	crgcggkggg	gacagagaca	gatttccgtg	ggacccgggt	2100
gggtgatgac	cgtagtcyga	gctgggcaga	gagggcgcgg	ggtcgtggac	atgaaacagg	2160
ccagcgagtr	gggacagcsg	gccaagaaac	cacctgcact	agggaggtgt	gagcatgggg	2220
acgagggcgg	ggcttgtgac	gagtgggcgg	ggccactgcc	gagacctggc	aggagcccaa	2280
tgggtgagcg	tggcgcattt	cccagctgga	atccggtgtc	gaagtggggg	cggggaccgc	2340
acctgtgctg	taagctcagt	gtgggtggcg	cggggcccgc	ggggtcttcc	ctgagtgcaa	2400
aggcggtcag	ggtgggcaga	gacgaggtgg	ggcaaagcct	gccccagcca	agggagcaag	2460
gtggatgcac	aaagagtggg	ccctgtgacc	agctggacag	agccagggac	tgcgggagac	2520
caggggagc	atagggttgg	agtgggtggt	ggatggtggg	gctaatgcct	tcatggccac	2580
gcgcacgtgc	ccgtcccacc	cccaggggtg	wtcctggcgc	gctatgggcc	cgcgyggcgc	2640
gagcagaggc	gcttctccrt	stccaccttg	cgcaacttgg	gcctgggcaa	gaagtcgctg	2700
gagsagtggg	tgaccragga	ggccgcctgc	ctttgtgccg	ccttcgccaa	ccactccggt	2760
gggtgatggg	cagaagggca	caaagcggga	actgggaagg	cgggggacgg	ggaaggcgac	2820
cccttacccg	catctcccac	ccccargacg	cccctttcgc	cccaacggtc	tcttggacaa	2880
agccgtgagc	aacgtgatcg	cctccctcac	ctgcgggcgc	cgcttcgagt	acgacgaccc	2940
tcgcttcctc	aggctgctgg	acctagctca	ggagggactg	aaggaggagt	cgggctttct	3000
gcgcgaggtg	cggagcgaga	gaccgaggag	tctctgcagg	gcgagctccc	gagaggtgcc	3060

SEQ Listing CYP2D6.ST25

ggggctggac tggggcctcg gaagagcagg atttgcatag atgggtttgg gaaaggacat 3120
tccaggagac cccactgtaa gaagggcctg gaggaggagg ggacatctca gacatggtcg 3180
tgggagaggt gtgcccgggt caggggcac caggagaggc caaggactct gtacctccta 3240
tccacgtcag agatttcgat tttaggtttc tcctctgggc aaggagagag grtggaggct 3300
ggcacttggg gagggacttg gtgaggtcag tggtaaggac aggcaggccc tgggtctacc 3360
tggagatggc tggggcctga gacttgtcca ggtgaacgca gagcacagga gggattgaga 3420
ccccgttctg tctggtgtag gtgctgaatg ctgtccccgt cctcctgcay atcccagcgc 3480
tggctggcaa ggtcctacgc ttccaaaagg ctttcctgac ccagctggat gagctgctaa 3540
ctgagcacag gatgacctgg gacccagccc agccccccg agacctgact gaggccttcc 3600
tggcagagat ggagaaggtg agagtggctg ccacggtggg gggcaagggt ggtgggttga 3660
gcgtcccagg aggaatgagg ggaggctggg caaaaggttg gaccagtgca tcacccggcg 3720
agccgcatct gggctgacag gtgcagaatt ggaggtcatt tgggggctac cccgttctgt 3780
cccgagtatg ctctcggccc tgctcaggcc aaggggaacc ctgagagcag cttcaatgat 3840
gagaacctgc gcatagtggt ggctgacctg ttctctgccg ggatggtgac cacctcgacc 3900
acgctggcct ggggcctcct gctcatgatc ctacatccgg atgtgcagcg tgagcccatc 3960
tgggaaacag tgcaggggcc gagggaggaa gggtacaggc gggggcccat gaactttgct 4020
gggacacccg gggctccaag cacaggcttg accaggatcc tgtaagcctg acctcctcca 4080
acataggagg caagaaggag tgtcagggcc ggacccctg ggtgctgacc cattgtgggg 4140
acgcatgtct gtccaggccg tgtccaacag gagatcgacg acrtgatagg gcaggtgcgg 4200
ygaccagaga tgggtgacca ggctcacatg ccctacacca ctgccgtgat tcaygaggtg 4260
cagcgctttg gggacatcgt ccccctgggt gtgacccata tgacatcccg tgacatcgaa 4320
gtacagggct tccgcatccc taaggtaggc ctggcgccct cctcacccca gctcagcacc 4380
agcmcctggt gatagcccca gcatggctac tgccaggtgg gcccactcta ggaamcctgg 4440
ccacctagtc ctcaatgcca ccacactgac tgtccccact tgggtggggg gtccagagta 4500
taggcagggc tggcctgtcc atccagagcc cccgtctagt ggggagacaa accaggacct 4560
gccagaatgt tggaggaccc aacgcctgca gggagagggg gcagtgtggg tgcctctgag 4620
aggtgtgact gcgccctgct gtggggtcgg agagggtact gtggagcttc tcgggcgcag 4680
gactagttga cagagtccag ctgtgtgcca ggcagtgtgt gtcccccgtg tgtttggtgg 4740
caggggtccc agcatcctag agtccagtcc ccactctcac cctgcatctc ctgcccaggg 4800
aacgacactc atcaccaacc tgtcatcggt gctgaaggat gaggccgtct gggagaagcc 4860
cttccgcttc caccccgaac acttcctgga tgcccagggc cactttgtga agccggaggc 4920
cttcctgcct ttctcagcag gtgcctgtgg ggagcccggc tccctgtccc cttccgtgga 4980
gtcttgcagg ggtatcaccc aggagccagg ctcactgacg cccctccccc ccccacaggc 5040
cgccgtgcat gcctcgggga gcccctggcc cgcatggagc tcttcctctt cttcacctcc 5100
ctgctgcagc acttcagctt ctcggtgccc actggacagc cccggcccag ccaccatggt 5160
gtctttgctt tcctggtgas cccatccccc tatgagcttt gtgctgtgcc cygctagaat 5220
ggggtaccta gtccccagcc tgctccctag ccagaggctc taatgtacaa taaagcaatg 5280
tggtagttcc aactcgggtc ccctgctcac gccctcgttg ggatcatcct cctcagggca 5340
accccacccc tgcctcattc ctgcttaccc caccgcctgg ccgcatttga gacaggggta 5400
cgttgaggct gagcagatgt cagttaccct tgcccataat cccatgtccc ccactgaccc 5460
aactctgact gcccagattg gtgacaagga ctacattgtc ctggcatgtg gggaaggggc 5520
cagaatgggc tgactagagg tgtcagtcag ccctggatgt ggtggagagg gcaggactca 5580
gcctggaggc ccatatttca ggcctaactc agcccacccc acatcaggga cagcagtcct 5640
gccagcacca tcacaacagt cacctccctt catatatgac accccaaaac ggaagacaaa 5700
tcatggcgtc agggagctat atgccagggc tacctacctc ccagggctca gtcggcaggt 5760
gccagaacgt tccctgggaa ggccccatgg aagcccagga ctgagccacc accctcagcc 5820
tcgtcacctc accacaggac tggctacctc tctgggccct cagggatgct gctgtacaga 5880
cccctgacca gtgacgagtt cgcactcagg gccaggctgg cgctggagga ggacacttgt 5940
ttggctccaa ccctaggtac catcctccca gtagggatca ggcagggccc acaggcctgc 6000
cctagggaca ggagtcaacc ttggacccat aaggcactgg ggcgggcaga gaaggaggag 6060
gtggcatggg cagctgagag ccagagaccc tgaccctagt ccttgctctg ccattacccc 6120
gtgtgacccc gggcccaccc ttccccaccc ttccccaccc cgggcttctg tttccttctg 6180
ccaacgagaa ggctgcttca cctgccccga gtcctgtctt cctgctctgc cttctggggc 6240
tgtggccctt gctggcctgg agccccaacc aagggcaggg actgctgtcc tccacgtctg 6300
tcctcaccga cataatgggc tgggctgggc acacaggcag tgcccaagag tttctaatga 6360
gcatatgatt acctgagtcc tgggcagacc ttcttaggga acagcctggg acagagaacc 6420
acagacactc tgaggagcca ccctgaggcc tcttttgcca gaggacccta ca 6472

<210> 2
<211> 1494
<212> DNA
<213> Homo sapiens

<400> 2
atggggctag aagcactggt gcccctggcc gtgatagtgg ccatcttcct gctcctggtg 60
gacctgatgc accggcgcca acgctgggct gcacgctacc caccaggccc cctgccactg 120
cccgggctgg gcaacctgct gcatgtggac ttccagaaca caccatactg cttcgaccag 180

SEQ Listing CYP2D6.ST25

ttgcggcgcc gcttcgggga cgtgttcagc ctgcagctgg cctggacgcc ggtggtcgtg 240
ctcaatgggc tggcggccgt gcgcgaggcg ctggtgaccc acggcgagga caccgccgac 300
cgcccgcctg tgcccatcac ccagatcctg ggtttcgggc cgcgttccca aggggtgttc 360
ctggcgcgct atgggcccgc gtggcgcgag cagaggcgct tctccgtgtc caccttgcgc 420
aacttgggcc tgggcaagaa gtcgctggag cagtgggtga ccgaggaggc cgcctgcctt 480
tgtgccgcct tcgccaacca ctccggacgc ccctttcgcc ccaacggtct cttggacaaa 540
gccgtgagca acgtgatcgc ctccctcacc tgcgggcgcc gcttcgagta cgacgaccct 600
cgcttcctca ggctgctgga cctagctcag gagggactga aggaggagtc gggctttctg 660
cgcgaggtgc tgaatgctgt ccccgtcctc ctgcatatcc cagcgctggc tggcaaggtc 720
ctacgcttcc aaaaggcttt cctgacccag ctggatgagc tgctaactga gcacaggatg 780
acctgggacc cagcccagcc ccccgagac ctgactgagg ccttcctggc agagatggag 840
aaggccaagg ggaaccctga gagcagcttc aatgatgaga acctgcgcat agtggtggct 900
gacctgttct ctgccgggat ggtgaccacc tcgaccacgc tggcctgggg cctcctgctc 960
atgatcctac atccggatgt gcagcgccgt gtccaacagg agatcgacga cgtgataggg 1020
caggtgcggc gaccagagat gggtgaccag gctcacatgc cctacaccac tgccgtgatt 1080
catgaggtgc agcgctttgg ggacatcgtc cccctgggtg tgacccatat gacatcccgt 1140
gacatcgaag tacagggctt ccgcatccct aagggaacga cactcatcac caacctgtca 1200
tcggtgctga aggatgaggc cgtctgggag aagcccttcc gcttccaccc cgaacacttc 1260
ctggatgccc agggccactt tgtgaagccg gaggccttcc tgcctttctc agcaggccgc 1320
cgtgcatgcc tcggggagcc cctggcccgc atggagctct tcctcttctt cacctccctg 1380
ctgcagcact tcagcttctc ggtgcccact ggacagcccc ggcccagcca ccatggtgtc 1440
tttgctttcc tggtgagccc atcccctat gagctttgtg ctgtgccccg ctag 1494

<210> 3
<211> 497
<212> PRT
<213> Homo sapiens

<400> 3

Met Gly Leu Glu Ala Leu Val Pro Leu Ala Val Ile Val Ala Ile Phe
1 5 10 15

Leu Leu Leu Val Asp Leu Met His Arg Arg Gln Arg Trp Ala Ala Arg
20 25 30

Tyr Pro Pro Gly Pro Leu Pro Leu Pro Gly Leu Gly Asn Leu Leu His
35 40 45

Val Asp Phe Gln Asn Thr Pro Tyr Cys Phe Asp Gln Leu Arg Arg Arg
50 55 60

Phe Gly Asp Val Phe Ser Leu Gln Leu Ala Trp Thr Pro Val Val Val
65 70 75 80

Leu Asn Gly Leu Ala Ala Val Arg Glu Ala Leu Val Thr His Gly Glu
85 90 95

Asp Thr Ala Asp Arg Pro Pro Val Pro Ile Thr Gln Ile Leu Gly Phe
100 105 110

Gly Pro Arg Ser Gln Gly Val Phe Leu Ala Arg Tyr Gly Pro Ala Trp
115 120 125

Arg Glu Gln Arg Arg Phe Ser Val Ser Thr Leu Arg Asn Leu Gly Leu
130 135 140

Gly Lys Lys Ser Leu Glu Gln Trp Val Thr Glu Glu Ala Ala Cys Leu
145 150 155 160

Cys Ala Ala Phe Ala Asn His Ser Gly Arg Pro Phe Arg Pro Asn Gly
165 170 175

Leu Leu Asp Lys Ala Val Ser Asn Val Ile Ala Ser Leu Thr Cys Gly
180 185 190

Arg Arg Phe Glu Tyr Asp Asp Pro Arg Phe Leu Arg Leu Leu Asp Leu

SEQ Listing CYP2D6.ST25

195 200 205

Ala Gln Glu Gly Leu Lys Glu Glu Ser Gly Phe Leu Arg Glu Val Leu
210 215 220

Asn Ala Val Pro Val Leu Leu His Ile Pro Ala Leu Ala Gly Lys Val
225 230 235 240

Leu Arg Phe Gln Lys Ala Phe Leu Thr Gln Leu Asp Glu Leu Leu Thr
245 250 255

Glu His Arg Met Thr Trp Asp Pro Ala Gln Pro Pro Arg Asp Leu Thr
260 265 270

Glu Ala Phe Leu Ala Glu Met Glu Lys Ala Lys Gly Asn Pro Glu Ser
275 280 285

Ser Phe Asn Asp Glu Asn Leu Arg Ile Val Val Ala Asp Leu Phe Ser
290 295 300

Ala Gly Met Val Thr Thr Ser Thr Thr Leu Ala Trp Gly Leu Leu Leu
305 310 315 320

Met Ile Leu His Pro Asp Val Gln Arg Arg Val Gln Gln Glu Ile Asp
325 330 335

Asp Val Ile Gly Gln Val Arg Arg Pro Glu Met Gly Asp Gln Ala His
340 345 350

Met Pro Tyr Thr Thr Ala Val Ile His Glu Val Gln Arg Phe Gly Asp
355 360 365

Ile Val Pro Leu Gly Val Thr His Met Thr Ser Arg Asp Ile Glu Val
370 375 380

Gln Gly Phe Arg Ile Pro Lys Gly Thr Thr Leu Ile Thr Asn Leu Ser
385 390 395 400

Ser Val Leu Lys Asp Glu Ala Val Trp Glu Lys Pro Phe Arg Phe His
405 410 415

Pro Glu His Phe Leu Asp Ala Gln Gly His Phe Val Lys Pro Glu Ala
420 425 430

Phe Leu Pro Phe Ser Ala Gly Arg Arg Ala Cys Leu Gly Glu Pro Leu
435 440 445

Ala Arg Met Glu Leu Phe Leu Phe Phe Thr Ser Leu Leu Gln His Phe
450 455 460

Ser Phe Ser Val Pro Thr Gly Gln Pro Arg Pro Ser His His Gly Val
465 470 475 480

Phe Ala Phe Leu Val Ser Pro Ser Pro Tyr Glu Leu Cys Ala Val Pro
485 490 495

Arg

<210> 4
<211> 15
<212> DNA
<213> Homo sapiens

<400> 4
gatggccrgg tccac 15

SEQ Listing CYP2D6.ST25

<210> 5
<211> 15
<212> DNA
<213> Homo sapiens

<400> 5
tgcaggcytc aggag 15

<210> 6
<211> 15
<212> DNA
<213> Homo sapiens

<400> 6
gaagcagsgg caaga 15

<210> 7
<211> 15
<212> DNA
<213> Homo sapiens

<400> 7
gggcaagrac ctctg 15

<210> 8
<211> 15
<212> DNA
<213> Homo sapiens

<400> 8
ggtgtgcyga gagtg 15

<210> 9
<211> 15
<212> DNA
<213> Homo sapiens

<400> 9
ggactagsac ctgta 15

<210> 10
<211> 15
<212> DNA
<213> Homo sapiens

<400> 10
gccgtgcrcg aggcg 15

<210> 11
<211> 15
<212> DNA
<213> Homo sapiens

<400> 11
gtgacccrcg gcgag 15

<210> 12
<211> 15
<212> DNA
<213> Homo sapiens

SEQ Listing CYP2D6.ST25

<400> 12
aggacacsgc cgacc 15

<210> 13
<211> 15
<212> DNA
<213> Homo sapiens

<400> 13
ccgcctgygc ccatc 15

<210> 14
<211> 15
<212> DNA
<213> Homo sapiens

<400> 14
gcccatcwcc cagat 15

<210> 15
<211> 15
<212> DNA
<213> Homo sapiens

<400> 15
cccatcaycc agatc 15

<210> 16
<211> 15
<212> DNA
<213> Homo sapiens

<400> 16
cacccagrtc ctggg 15

<210> 17
<211> 15
<212> DNA
<213> Homo sapiens

<400> 17
tcctgggytt cgggc 15

<210> 18
<211> 15
<212> DNA
<213> Homo sapiens

<400> 18
ggcaagcrgc ggtgg 15

<210> 19
<211> 15
<212> DNA
<213> Homo sapiens

<400> 19
gcagcggkgg ggaca 15

<210> 20

SEQ Listing CYP2D6.ST25

<211> 15
<212> DNA
<213> Homo sapiens

<400> 20
cgtagtcyga gctgg 15

<210> 21
<211> 15
<212> DNA
<213> Homo sapiens

<400> 21
ggacagcsgg ccaag 15

<210> 22
<211> 15
<212> DNA
<213> Homo sapiens

<400> 22
aggggtgwtc ctggc 15

<210> 23
<211> 15
<212> DNA
<213> Homo sapiens

<400> 23
gcccgcgygg cgcga 15

<210> 24
<211> 15
<212> DNA
<213> Homo sapiens

<400> 24
cttctccrtg tccac 15

<210> 25
<211> 15
<212> DNA
<213> Homo sapiens

<400> 25
ggtgaccrag gaggc 15

<210> 26
<211> 15
<212> DNA
<213> Homo sapiens

<400> 26
agagaggrtg gaggc 15

<210> 27
<211> 15
<212> DNA
<213> Homo sapiens

<400> 27

cgacgacrtg atagg 15

<210> 28
<211> 15
<212> DNA
<213> Homo sapiens

<400> 28
ggtgcggyga ccaga 15

<210> 29
<211> 15
<212> DNA
<213> Homo sapiens

<400> 29
tgattcayga ggtgc 15

<210> 30
<211> 15
<212> DNA
<213> Homo sapiens

<400> 30
caccagcmcc tggtg 15

<210> 31
<211> 15
<212> DNA
<213> Homo sapiens

<400> 31
ctaggaamcc tggcc 15

<210> 32
<211> 15
<212> DNA
<213> Homo sapiens

<400> 32
tgtgcccygc tagaa 15

<210> 33
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 33
gaggtggatg gccrg 15

<210> 34
<211> 15
<212> DNA
<213> Homo sapiens

<400> 34
gtttcagtgg accyg 15

<210> 35
<211> 15

SEQ Listing CYP2D6.ST25

<212> DNA
<213> Homo Sapiens

<400> 35
aggctttgca ggcyt 15

<210> 36
<211> 15
<212> DNA
<213> Homo sapiens

<400> 36
tccaagctcc tgarg 15

<210> 37
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 37
gagcaggaag cagsg 15

<210> 38
<211> 15
<212> DNA
<213> Homo sapiens

<400> 38
agaggttctt gccsc 15

<210> 39
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 39
aagcaggggc aagra 15

<210> 40
<211> 15
<212> DNA
<213> Homo sapiens

<400> 40
ctgctccaga ggtyc 15

<210> 41
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 41
gcgctcggtg tgcyg 15

<210> 42
<211> 15
<212> DNA
<213> Homo sapiens

<400> 42
gcaggacact ctcrg 15

SEQ Listing CYP2D6.ST25

<210> 43
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 43
caaataggac tagsa 15

<210> 44
<211> 15
<212> DNA
<213> Homo sapiens

<400> 44
ccagactaca ggtsc 15

<210> 45
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 45
ctggcggccg tgcrc 15

<210> 46
<211> 15
<212> DNA
<213> Homo sapiens

<400> 46
caccagcgcc tcgyg 15

<210> 47
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 47
gcgctggtga cccrc 15

<210> 48
<211> 15
<212> DNA
<213> Homo sapiens

<400> 48
ggtgtcctcg ccgyg 15

<210> 49
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 49
acggcgagga cacsg 15

<210> 50
<211> 15
<212> DNA

SEQ Listing CYP2D6.ST25

<213> Homo sapiens

<400> 50
gcgggcggtc ggcsg 15

<210> 51
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 51
gaccgcccgc ctgyg 15

<210> 52
<211> 15
<212> DNA
<213> Homo sapiens

<400> 52
ctgggtgatg ggcrc 15

<210> 53
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 53
gcctgtgccc atcwc 15

<210> 54
<211> 15
<212> DNA
<213> Homo sapiens

<400> 54
cccaggatct gggwg 15

<210> 55
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 55
cctgtgccca tcayc 15

<210> 56
<211> 15
<212> DNA
<213> Homo sapiens

<400> 56
acccaggatc tggrt 15

<210> 57
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 57
gcccatcacc cagrt 15

SEQ Listing CYP2D6.ST25

<210> 58
<211> 15
<212> DNA
<213> Homo sapiens

<400> 58
ccgaaaccca ggayc 15

<210> 59
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 59
cccagatcct gggyt 15

<210> 60
<211> 15
<212> DNA
<213> Homo sapiens

<400> 60
aacgcggccc gaarc 15

<210> 61
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 61
tcccaaggca agcrg 15

<210> 62
<211> 15
<212> DNA
<213> Homo sapiens

<400> 62
ctgtccccac cgcyg 15

<210> 63
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 63
aggcaagcag cggkg 15

<210> 64
<211> 15
<212> DNA
<213> Homo sapiens

<400> 64
tgtctctgtc cccmc 15

<210> 65
<211> 15
<212> DNA
<213> Homo Sapiens

SEQ Listing CYP2D6.ST25

<400> 65
gatgaccgta gtcyg 15

<210> 66
<211> 15
<212> DNA
<213> Homo sapiens

<400> 66
ctctgcccag ctcrg 15

<210> 67
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 67
gagtggggac agcsg 15

<210> 68
<211> 15
<212> DNA
<213> Homo sapiens

<400> 68
tggtttcttg gccsg 15

<210> 69
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 69
acccccaggg gtgwt 15

<210> 70
<211> 15
<212> DNA
<213> Homo sapiens

<400> 70
tagcgcgcca ggawc 15

<210> 71
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 71
ctatgggccc gcgyg 15

<210> 72
<211> 15
<212> DNA
<213> Homo sapiens

<400> 72
ctctgctcgc gccrc 15

SEQ Listing CYP2D6.ST25

<210> 73
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 73
gaggcgcttc tccrt 15

<210> 74
<211> 15
<212> DNA
<213> Homo sapiens

<400> 74
cgcaaggtgg acayg 15

<210> 75
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 75
gcagtgggtg accra 15

<210> 76
<211> 15
<212> DNA
<213> Homo sapiens

<400> 76
caggcggcct cctyg 15

<210> 77
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 77
gcaaggagag aggrt 15

<210> 78
<211> 15
<212> DNA
<213> Homo sapiens

<400> 78
gtgccagcct ccayc 15

<210> 79
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 79
ggagatcgac gacrt 15

<210> 80
<211> 15
<212> DNA
<213> Homo sapiens

<400> 80
acctgcccta tcayg 15

<210> 81
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 81
agggcaggtg cggyg 15

<210> 82
<211> 15
<212> DNA
<213> Homo sapiens

<400> 82
cccatctctg gtcrc 15

<210> 83
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 83
ctgccgtgat tcayg 15

<210> 84
<211> 15
<212> DNA
<213> Homo sapiens

<400> 84
agcgctgcac ctcrt 15

<210> 85
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 85
gctcagcacc agcmc 15

<210> 86
<211> 15
<212> DNA
<213> Homo sapiens

<400> 86
ggctatcacc aggkg 15

<210> 87
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 87
cccactctag gaamc 15

<210> 88

<211> 15
<212> DNA
<213> Homo sapiens

<400> 88
ctaggtggcc aggkt 15

<210> 89
<211> 15
<212> DNA
<213> Homo Sapiens

<400> 89
ttgtgctgtg cccyg 15

<210> 90
<211> 15
<212> DNA
<213> Homo sapiens

<400> 90
accccattct agcrg 15

<210> 91
<211> 10
<212> DNA
<213> Homo sapiens

<400> 91
gtggatggcc 10

<210> 92
<211> 10
<212> DNA
<213> Homo sapiens

<400> 92
tcagtggacc 10

<210> 93
<211> 10
<212> DNA
<213> Homo sapiens

<400> 93
ctttgcaggc 10

<210> 94
<211> 10
<212> DNA
<213> Homo sapiens

<400> 94
aagctcctga 10

<210> 95
<211> 10
<212> DNA
<213> Homo sapiens

<400> 95

caggaagcag 10

<210> 96
<211> 10
<212> DNA
<213> Homo sapiens

<400> 96
ggttcttgcc 10

<210> 97
<211> 10
<212> DNA
<213> Homo sapiens

<400> 97
caggggcaag 10

<210> 98
<211> 10
<212> DNA
<213> Homo sapiens

<400> 98
ctccagaggt 10

<210> 99
<211> 10
<212> DNA
<213> Homo sapiens

<400> 99
ctcggtgtgc 10

<210> 100
<211> 10
<212> DNA
<213> Homo sapiens

<400> 100
ggacactctc 10

<210> 101
<211> 10
<212> DNA
<213> Homo sapiens

<400> 101
ataggactag 10

<210> 102
<211> 10
<212> DNA
<213> Homo sapiens

<400> 102
gactacaggt 10

<210> 103
<211> 10

SEQ Listing CYP2D6.ST25

<212> DNA
<213> Homo sapiens

<400> 103
gcggccgtgc 10

<210> 104
<211> 10
<212> DNA
<213> Homo sapiens

<400> 104
cagcgcctcg 10

<210> 105
<211> 10
<212> DNA
<213> Homo sapiens

<400> 105
ctggtgaccc 10

<210> 106
<211> 10
<212> DNA
<213> Homo sapiens

<400> 106
gtcctcgccg 10

<210> 107
<211> 10
<212> DNA
<213> Homo sapiens

<400> 107
gcgaggacac 10

<210> 108
<211> 10
<212> DNA
<213> Homo sapiens

<400> 108
ggcggtcggc 10

<210> 109
<211> 10
<212> DNA
<213> Homo sapiens

<400> 109
cgcccgcctg 10

<210> 110
<211> 10
<212> DNA
<213> Homo sapiens

<400> 110
ggtgatgggc 10

<210> 111
<211> 10
<212> DNA
<213> Homo sapiens

<400> 111
tgtgcccatc 10

<210> 112
<211> 10
<212> DNA
<213> Homo sapiens

<400> 112
aggatctggg 10

<210> 113
<211> 10
<212> DNA
<213> Homo sapiens

<400> 113
gtgcccatca 10

<210> 114
<211> 10
<212> DNA
<213> Homo sapiens

<400> 114
caggatctgg 10

<210> 115
<211> 10
<212> DNA
<213> Homo sapiens

<400> 115
catcacccag 10

<210> 116
<211> 10
<212> DNA
<213> Homo sapiens

<400> 116
aaacccagga 10

<210> 117
<211> 10
<212> DNA
<213> Homo sapiens

<400> 117
agatcctggg 10

<210> 118
<211> 10
<212> DNA

<213> Homo sapiens

<400> 118
gcggcccgaa 10

<210> 119
<211> 10
<212> DNA
<213> Homo sapiens

<400> 119
caaggcaagc 10

<210> 120
<211> 10
<212> DNA
<213> Homo sapiens

<400> 120
tccccaccgc 10

<210> 121
<211> 10
<212> DNA
<213> Homo sapiens

<400> 121
caagcagcgg 10

<210> 122
<211> 10
<212> DNA
<213> Homo sapiens

<400> 122
ctctgtcccc 10

<210> 123
<211> 10
<212> DNA
<213> Homo sapiens

<400> 123
gaccgtagtc 10

<210> 124
<211> 10
<212> DNA
<213> Homo sapiens

<400> 124
tgcccagctc 10

<210> 125
<211> 10
<212> DNA
<213> Homo sapiens

<400> 125
tggggacagc 10

SEQ Listing CYP2D6.ST25

<210> 126
<211> 10
<212> DNA
<213> Homo sapiens

<400> 126
tttcttggcc 10

<210> 127
<211> 10
<212> DNA
<213> Homo sapiens

<400> 127
cccaggggtg 10

<210> 128
<211> 10
<212> DNA
<213> Homo sapiens

<400> 128
cgcgccagga 10

<210> 129
<211> 10
<212> DNA
<213> Homo sapiens

<400> 129
tgggcccgcg 10

<210> 130
<211> 10
<212> DNA
<213> Homo sapiens

<400> 130
tgctcgcgcc 10

<210> 131
<211> 10
<212> DNA
<213> Homo sapiens

<400> 131
gcgcttctcc 10

<210> 132
<211> 10
<212> DNA
<213> Homo sapiens

<400> 132
aaggtggaca 10

<210> 133
<211> 10
<212> DNA
<213> Homo sapiens

SEQ Listing CYP2D6.ST25

<400> 133
gtgggtgacc 10

<210> 134
<211> 10
<212> DNA
<213> Homo sapiens

<400> 134
gcggcctcct 10

<210> 135
<211> 10
<212> DNA
<213> Homo sapiens

<400> 135
aggagagagg 10

<210> 136
<211> 10
<212> DNA
<213> Homo sapiens

<400> 136
ccagcctcca 10

<210> 137
<211> 10
<212> DNA
<213> Homo sapiens

<400> 137
gatcgacgac 10

<210> 138
<211> 10
<212> DNA
<213> Homo sapiens

<400> 138
tgccctatca 10

<210> 139
<211> 10
<212> DNA
<213> Homo sapiens

<400> 139
gcaggtgcgg 10

<210> 140
<211> 10
<212> DNA
<213> Homo sapiens

<400> 140
atctctggtc 10

SEQ Listing CYP2D6.ST25

<210> 141
<211> 10
<212> DNA
<213> Homo sapiens

<400> 141
ccgtgattca 10

<210> 142
<211> 10
<212> DNA
<213> Homo sapiens

<400> 142
gctgcacctc 10

<210> 143
<211> 10
<212> DNA
<213> Homo sapiens

<400> 143
cagcaccagc 10

<210> 144
<211> 10
<212> DNA
<213> Homo sapiens

<400> 144
tatcaccagg 10

<210> 145
<211> 10
<212> DNA
<213> Homo sapiens

<400> 145
actctaggaa 10

<210> 146
<211> 10
<212> DNA
<213> Homo sapiens

<400> 146
ggtggccagg 10

<210> 147
<211> 10
<212> DNA
<213> Homo sapiens

<400> 147
tgctgtgccc 10

<210> 148
<211> 10
<212> DNA
<213> Homo sapiens

<400> 148
ccattctagc 10

<210> 149
<211> 18
<212> DNA
<213> Homo sapiens

<400> 149
tgtaaaacga cggccagt 18

<210> 150
<211> 19
<212> DNA
<213> Homo sapiens

<400> 150
aggaaacagc tatgaccat 19

<210> 151
<211> 5040
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (30)..(30)
<223> PS1: polymorphic base G or A

<220>
<221> misc_feature
<222> (61)..(120)
<223> n's represent sequence between PS1and PS2

<220>
<221> allele
<222> (150)..(150)
<223> PS2: polymorphic base T or C

<220>
<221> misc_feature
<222> (181)..(240)
<223> n's represent sequence between PS2 and PS3

<220>
<221> allele
<222> (270)..(270)
<223> PS3 : polymorphic base G or C

<220>
<221> misc_feature
<222> (301)..(360)
<223> n's represent sequence between PS3 and PS4

<220>
<221> allele
<222> (390)..(390)
<223> polymorphic base A or G

SEQ Listing CYP2D6.ST25

<220>
<221> misc_feature
<222> (421)..(480)
<223> n's represent sequence between PS4 and PS5

<220>
<221> allele
<222> (510)..(510)
<223> PS5: polymorphic base G or A

<220>
<221> misc_feature
<222> (541)..(600)
<223> n's represent sequence between PS5 and PS6

<220>
<221> allele
<222> (630)..(630)
<223> PS6: polymorphic base T or C

<220>
<221> misc_feature
<222> (661)..(720)
<223> n's represent sequence between PS6 and PS7

<220>
<221> allele
<222> (750)..(750)
<223> PS7: polymorphic base G or A

<220>
<221> misc_feature
<222> (781)..(840)
<223> n's represent sequence between PS7 and PS8

<220>
<221> allele
<222> (870)..(870)
<223> PS8: polymorphic base G or A

<220>
<221> misc_feature
<222> (901)..(960)
<223> n's represent sequence between PS8 and PS9

<220>
<221> allele
<222> (990)..(990)
<223> PS9: polymorphic base C or T

<220>
<221> misc_feature
<222> (1021)..(1080)
<223> n's represent sequence between PS9 and PS10

<220>

SEQ Listing CYP2D6.ST25

<221> allele
<222> (1110)..(1110)
<223> PS10: polymorphic base G or C

<220>
<221> misc_feature
<222> (1141)..(1200)
<223> n's represent sequence between PS10 and PS11

<220>
<221> allele
<222> (1230)..(1230)
<223> PS11: polymorphic base T or G

<220>
<221> misc_feature
<222> (1261)..(1320)
<223> n's represent sequence between PS11 and PS12

<220>
<221> allele
<222> (1350)..(1350)
<223> PS12: polymorphic base G or A

<220>
<221> misc_feature
<222> (1381)..(1440)
<223> n's represent sequence between PS12 and PS13

<220>
<221> allele
<222> (1470)..(1470)
<223> PS13: polymorphic base C or A

<220>
<221> misc_feature
<222> (1501)..(1560)
<223> n's represent sequence between PS13 and PS14

<220>
<221> allele
<222> (1590)..(1590)
<223> PS14: polymorphic base A or G

<220>
<221> misc_feature
<222> (1621)..(1680)
<223> n's represent sequence between PS14 and PS15

<220>
<221> allele
<222> (1710)..(1710)
<223> PS15: polymorphic base C or G

<220>
<221> misc_feature
<222> (1741)..(1800)

SEQ Listing CYP2D6.ST25

<223> n's represent sequence between PS15 and PS16

<220>
<221> allele
<222> (1830)..(1830)
<223> PS16: polymorphic base T or C

<220>
<221> misc_feature
<222> (1861)..(1920)
<223> n's represent sequence between PS16 and PS17

<220>
<221> allele
<222> (1950)..(1950)
<223> PS17: polymorphic base A or T

<220>
<221> misc_feature
<222> (1981)..(2040)
<223> n's represent sequence between PS17 and PS18

<220>
<221> allele
<222> (2070)..(2070)
<223> PS18: polymorphic base C or T

<220>
<221> misc_feature
<222> (2101)..(2160)
<223> n's represent sequence between PS18 and PS19

<220>
<221> allele
<222> (2190)..(2190)
<223> PS19: polymorphic base A or G

<220>
<221> misc_feature
<222> (2221)..(2280)
<223> n's represent sequence between PS19 and PS20

<220>
<221> allele
<222> (2310)..(2310)
<223> PS20: polymorphic base T or C

<220>
<221> misc_feature
<222> (2341)..(2400)
<223> n's represent sequence between PS20 and PS21

<220>
<221> allele
<222> (2430)..(2430)
<223> PS21: polymorphic base C or T

SEQ Listing CYP2D6.ST25

<220>
<221> misc_feature
<222> (2461)..(2520)
<223> n's represent sequence between PS21and PS22

<220>
<221> allele
<222> (2550)..(2550)
<223> PS22: polymorphic base A or G

<220>
<221> misc_feature
<222> (2581)..(2640)
<223> n's represent sequence between PS22 and PS23

<220>
<221> allele
<222> (2670)..(2670)
<223> PS23: polymorphic base T or G

<220>
<221> misc_feature
<222> (2701)..(2760)
<223> n's represent sequence between PS23 and PS24

<220>
<221> allele
<222> (2790)..(2790)
<223> PS24: polymorphic base C or T

<220>
<221> misc_feature
<222> (2821)..(2880)
<223> n's represent sequence between PS24 and PS25

<220>
<221> allele
<222> (2910)..(2910)
<223> PS25: polymorphic base G or A

<220>
<221> misc_feature
<222> (2941)..(3000)
<223> n's represent sequence between PS25 and PS26

<220>
<221> allele
<222> (3030)..(3030)
<223> PS26: polymorphic base G or C

<220>
<221> misc_feature
<222> (3061)..(3120)
<223> n's represent sequence between PS26 and PS27

<220>

SEQ Listing CYP2D6.ST25

<221> allele
<222> (3150)..(3150)
<223> PS27: polymorphic base T or A

<220>
<221> misc_feature
<222> (3181)..(3240)
<223> n's represent sequence between PS27and PS28

<220>
<221> allele
<222> (3270)..(3270)
<223> PS28: polymorphic base T or C

<220>
<221> misc_feature
<222> (3301)..(3360)
<223> n's represent sequence between PS28 and PS29

<220>
<221> allele
<222> (3390)..(3390)
<223> PS29: polymorphic base G or A

<220>
<221> misc_feature
<222> (3421)..(3480)
<223> n's represent sequence between PS29 and PS30

<220>
<221> allele
<222> (3510)..(3510)
<223> PS30: polymorphic base G or C

<220>
<221> misc_feature
<222> (3541)..(3600)
<223> n's represent sequence between PS30 and PS31

<220>
<221> allele
<222> (3630)..(3630)
<223> PS31: polymorphic base C or G

<220>
<221> misc_feature
<222> (3661)..(3720)
<223> n's represent sequence between PS31 and PS32

<220>
<221> allele
<222> (3750)..(3750)
<223> PS32: polymorphic base G or A

<220>
<221> misc_feature
<222> (3781)..(3840)

SEQ Listing CYP2D6.ST25

<223> n's represent sequence between PS32 and PS33

<220>
<221> allele
<222> (3870)..(3870)
<223> PS33: polymorphic base G or A

<220>
<221> misc_feature
<222> (3901)..(3960)
<223> n's represent sequence between PS33and PS34

<220>
<221> allele
<222> (3990)..(3990)
<223> PS34: polymorphic base G or A

<220>
<221> misc_feature
<222> (4021)..(4080)
<223> n's represent sequence between PS34 and PS35

<220>
<221> allele
<222> (4110)..(4110)
<223> PS35: polymorphic base T or C

<220>
<221> misc_feature
<222> (4141)..(4200)
<223> n's represent sequence between PS35 and PS36

<220>
<221> allele
<222> (4230)..(4230)
<223> PS36: polymorphic base G or A

<220>
<221> misc_feature
<222> (4261)..(4320)
<223> n's represent sequence between PS36 and PS37

<220>
<221> allele
<222> (4350)..(4350)
<223> PS37: polymorphic base C or T

<220>
<221> misc_feature
<222> (4381)..(4440)
<223> n's represent sequence between PS37 and PS38

<220>
<221> allele
<222> (4470)..(4470)
<223> PS38: polymorphic base T or C

SEQ Listing CYP2D6.ST25

<220>
<221> misc_feature
<222> (4501)..(4560)
<223> n's represent sequence between PS38 and PS39

<220>
<221> allele
<222> (4590)..(4590)
<223> PS39: polymorphic base A or C

<220>
<221> misc_feature
<222> (4621)..(4680)
<223> n's represent sequence between PS39 and PS40

<220>
<221> allele
<222> (4710)..(4710)
<223> PS40: polymorphic base C or A

<220>
<221> misc_feature
<222> (4741)..(4800)
<223> n's represent sequence between PS40 and PS41

<220>
<221> allele
<222> (4830)..(4830)
<223> PS41: polymorphic base G or C

<220>
<221> misc_feature
<222> (4861)..(4920)
<223> n's represent sequence between PS41 and PS42

<220>
<221> allele
<222> (4950)..(4950)
<223> PS42: polymorphic base C or T

<220>
<221> misc_feature
<222> (4981)..(5040)
<223> n's represent sequence 3' to PS42

<400> 151
gtatgagcct agctgggagg tggatggccr ggtccactga aaccctggtt atcccagaag 60
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 120
ccctggttat cccagaaggc tttgcaggcy tcaggagctt ggagtgggga gagggggtga 180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240
cctgggtaag ggcctggagc aggaagcags ggcaagaacc tctggagcag cccatacccg 300
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 360
aagggcctgg agcaggaagc aggggcaagr acctctggag cagcccatac ccgccctggc 420
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 480
cccgccctgg cctgactctg ccactggcar cacagtcaac acagcaggtt cactcacagc 540
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 600
ataagggaag ggtcacgcgc tcggtgtgcy gagagtgtcc tgcctggtcc tctgtgcctg 660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 720

SEQ Listing CYP2D6.ST25

gtgaggcagg tatggggcta gaagcactgr tgcccctggc cgtgatagtg gccatcttcc 780
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 840
tggggctaga agcactggtg cccctggccr tgatagtggc catcttcctg ctcctggtgg 900
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 960
accggcgcca acgctgggct gcacgctacy caccaggccc cctgccactg cccgggctgg 1020
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1080
ggggtcgtcc aaggttcaaa taggactags acctgtagtc tggggtgatc ctggcttgac 1140
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1200
caaataggac taggacctgt agtctggggk gatcctggct tgacaagagg ccctgaccct 1260
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1320
ggtcgtgctc aatgggctgg cggccgtgcr cgaggcgctg gtgacccacg gcgaggacac 1380
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1440
tcaatgggct ggcggccgtg cgcgaggcgm tggtgaccca cggcgaggac accgccgacc 1500
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1560
ggcggccgtg cgcgaggcgc tggtgacccr cggcgaggac accgccgacc gcccgcctgt 1620
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1680
gaggcgctgg tgacccacgg cgaggacacs gccgaccgcc cgcctgtgcc catcacccag 1740
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1800
cggcgaggac accgccgacc gcccgcctgy gcccatcacc cagatcctgg gtttcgggcc 1860
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1920
acaccgccga ccgcccgcct gtgcccatcw cccagatcct gggtttcggg ccgcgttccc 1980
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2040
caccgccgac cgcccgcctg tgcccatcay ccagatcctg ggtttcgggc cgcgttccca 2100
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2160
ccgaccgccc gcctgtgccc atcacccagr tcctgggttt cgggccgcgt tcccaaggca 2220
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2280
ccgcctgtgc ccatcaccca gatcctgggy ttcgggccgc gttcccaagg caagcagcgg 2340
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2400
cctgtgccca tcacccagat cctgggttty gggccgcgtt cccaaggcaa gcagcggtgg 2460
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2520
gggtttcggg ccgcgttccc aaggcaagcr gcggtgggga cagagacaga tttccgtggg 2580
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2640
tcgggccgcg ttcccaaggc aagcagcggk ggggacagag acagatttcc gtgggacccg 2700
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2760
tgggacccgg gtgggtgatg accgtagtcy gagctgggca gagagggcgc ggggtcgtgg 2820
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 2880
ggtcgtggac atgaaacagg ccagcgagtr gggacagcgg gccaagaaac cacctgcact 2940
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3000
catgaaacag gccagcgagt ggggacagcs ggccaagaaa ccacctgcac tagggaggtg 3060
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3120
cgcacgtgcc cgtcccaccc ccaggggtgw tcctggcgcg ctatgggccc gcgtggcgcg 3180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3240
gggtgttcct ggcgcgctat gggcccgcgy ggcgcgagca gaggcgcttc tccgtgtcca 3300
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3360
ccgcgtggcg cgagcagagg cgcttctccr tgtccacctt gcgcaacttg ggcctgggca 3420
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3480
gcgtggcgcg agcagaggcg cttctccgts tccaccttgc gcaacttggg cctgggcaag 3540
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3600
acttgggcct gggcaagaag tcgctggags agtgggtgac cgaggaggcc gcctgccttt 3660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3720
gcaagaagtc gctggagcag tgggtgaccr aggaggccgc ctgcctttgt gccgccttcg 3780
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3840
cgacccctta cccgcatctc ccacccccar gacgcccctt tcgccccaac ggtctcttgg 3900
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3960
taggtttctc ctctgggcaa ggagagaggr tggaggctgg cacttgggga gggacttggt 4020
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4080
gtgctgaatg ctgtccccgt cctcctgcay atcccagcgc tggctggcaa ggtcctacgc 4140
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4200
caggccgtgt ccaacaggag atcgacgacr tgatagggca ggtgcggcga ccagagatgg 4260
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4320
agatcgacga cgtgataggg caggtgcggy gaccagagat gggtgaccag gctcacatgc 4380
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4440
cacatgccct acaccactgc cgtgattcay gaggtgcagc gctttgggga catcgtcccc 4500
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4560
cgccctcctc accccagctc agcaccagcm cctggtgata gccccagcat ggctactgcc 4620
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4680
gctactgcca ggtgggccca ctctaggaam cctggccacc tagtcctcaa tgccaccaca 4740
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4800

SEQ Listing CYP2D6.ST25

ccaccatggt gtctttgctt tcctggtgas cccatccccc tatgagcttt gtgctgtgcc 4860
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4920
catcccccta tgagctttgt gctgtgcccy gctagaatgg ggtacctagt ccccagcctg 4980
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5040

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 May 2002 (16.05.2002)

PCT

(10) International Publication Number
WO 02/038589 A3

(51) International Patent Classification[7]: C12Q 1/68, C07H 21/04, C12N 15/12, C07K 14/00, 16/18, G01N 33/48

(21) International Application Number: PCT/US01/47396

(22) International Filing Date:
9 November 2001 (09.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/247,943 9 November 2000 (09.11.2000) US

(71) Applicant *(for all designated States except US)*: **GENAISSANCE PHARMACEUTICALS, INC.** [US/US]; Five Science Park, New Haven, CT 06511 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ANASTASIO, Alison, E.** [US/US]; 367 Alden Avenue #8B, New Haven, CT 06515 (US). **CHEW, Anne** [US/US]; 1477 Beacon Street, #64, Brookline, MA 02446 (US). **CHOI, Julie, Y.** [US/US]; 38 Elizabeth Street, West Haven, CT 06516 (US). **DENTON, R., Rex** [US/US]; 129 Hunter's Trail, Madison, CT 06443 (US). **NANDABALAN, Krishnan** [US/US]; 228 Village Pond Road, Guilford, CT 06437 (US). **PETERSEN, Nathan** [US/US]; 300 Eastern Street, Apartment 288, New Haven, CT 06513 (US). **ROUNDS, Eileen** [US/US]; 36 Patton Drive, Cheshire, CT 06410 (US).

(74) Agents: **SHTIVELBAND, Inna** et al.; Genaissance Pharmaceuticals, Inc., Five Science Park, New Haven, CT 06511 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
25 July 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/038589 A3

(54) Title: HAPLOTYPES OF THE CYP2D6 GENE

(57) Abstract: Novel genetic variants of the Cytochrome P450, Subfamily IID, Polypeptide 6 (CYP2D6) gene are described. Various genotypes, haplotypes, and haplotype pairs that exist in the general United States population are disclosed for the CYP2D6 gene. Compositions and methods for haplotyping and/or genotyping the CYP2D6 gene in an individual are also disclosed. Polynucleotides defined by the haplotypes disclosed herein are also described

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/47396

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12Q 1/68; C07H 21/04; C12N 15/12; C07K 14/00, 16/18; G01N 33/48
US CL : 435/6, 7.1; 536/23.1, 23.2, 23.5; 800/8; 530/350, 387.1; 702/20

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/6, 7.1; 536/23.1, 23.2, 23.5; 800/8; 530/350, 387.1; 702/20

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MCLELLAN, R.A. et al. Characterization and functional analysis of two common human cytochrome P450 1B1 variants. Archives of Biochemistry and Biophysics. 01 June 2000, Vol. 378, No. 1, pages 175-181, see entire document.	1-34
Y	MUNOZ, S. et al. Genetic polymorphisms of CYP2D6, CYP1A1 and CYP2E1 in the South-Amerindian population of Chile. Pharmacogenetics. 1998, Vol. 8, pages 343-351, see entire document.	1-34
Y	NISHIDA, Y. et al, CYP2D6 genotypes in a Japanese population: low frequencies of CYP2D6 gene duplication but high frequency of CYP2D6*10. Pharmacogenetics. 2000, Vol. 10, pages 567-570, see entire document.	1-34
Y	US 5,965,396 A (PAN et al) 12 October 1999 (12.10.1999), see especially column 6, line 17 through column 25, line 61.	1-34
Y	OUELLETTE, B.F.F. et al. The GenBank sequence database. In: Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. (Baxevanis et al, Eds.), 1998, Wiely-Liss, Inc., (New York, NY, USA), pages 16-45, see entire document	33

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 March 2002 (28.03.2002)	29 APR 2002
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer James Martinell Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/47396

Continuation of B. FIELDS SEARCHED Item 3:
WEST, STN (CAPLUS)
search terms: cytochrome P450, single nucleotide polymorphism, SNP, CYP2D6, haplotype